



TITLE:

Structure and Gene Expression of Rice Mitochondrial Genome(Dissertation_全文)

AUTHOR(S):

Yamato, Katsuyuki

CITATION:

Yamato, Katsuyuki. Structure and Gene Expression of Rice Mitochondrial Genome. 京都大学, 1993, 博士(農学)

ISSUE DATE:

1993-03-23

URL:

<https://doi.org/10.11501/3066277>

RIGHT:

新 制
農
649
京大附図

**Structure and Gene Expression
of
Rice Mitochondrial Genome**

Katsuyuki Yamato
1993

**Structure and Gene Expression
of
Rice Mitochondrial Genome**

**Katsuyuki Yamato
1993**

...Quam magnificata sunt opera tua.

Omnia in sapientia fecisti;

impleta est terra possessione tua.

Contents

Introduction	1
1 Physical Structure of Rice Mitochondrial Genome	5
2 Structure and Expression of the Three-copied <i>rrn18</i> Gene	27
3 The Evolution of Rice	47
Summary	55
References	57
List of Publications	64
Acknowledgements	66

Abbreviations

AMV	avian myeloblastosis virus
ATP	adenosine triphosphate
bp	base pair(s)
cpm	count per minute
CTAB	cetyltrimethylammonium bromide
dATP	deoxyadenosine triphosphate
dCTP	deoxycytidine triphosphate
dGTP	deoxyguanosine triphosphate
DNA	deoxyribonucleic acid
DTT	dithiothreitol
dTTP	deoxythymidine triphosphate
EDTA	ethylenediamine tetraacetic acid
GTP	guanosine triphosphate
kb	kilobase pair(s)
nt	nucleotide(s)
PIPES	piperazine- <i>N,N'</i> -bis-(2-ethanesulfonic acid)
RNA	ribonucleic acid
rRNA	ribosomal RNA
SDS	sodium dodecyl sulfate
SSC	standard saline citrate
tRNA	transfer RNA

E. coli *Escherichia coli*

Introduction

Plant mitochondrial genomes have been intensively investigated at the molecular level in order to clarify the mechanism of cytoplasmic male sterility (CMS), since several observations to date have indicated that a major cytoplasmic factor of the CMS trait resides in mitochondria rather than in chloroplasts. Certain remarkable findings during the process of these investigations, however, have indicated that the plant mitochondrial genetic system is itself one of the most fascinating subjects in molecular biology.

Plant mitochondrial genomes are the largest mitochondrial genomes of any other organisms to have been studied to date. In addition, heterogeneous molecular species are generally found within plant mitochondria. The heterogeneity of plant mitochondrial genomes arises from their multipartite structure, as well as the presence of several kinds of plasmids. These characteristics make it difficult to determine both the size and overall gene organization of plant mitochondrial genome.

Although the mitochondrial genome has a heterogeneous structure, its physical map has, nevertheless, been constructed for a number of plant species (Palmer and Shields 1984; Lonsdale *et al.* 1984; Chétrit *et al.* 1984; Stern and Palmer 1986; Palmer and Herbon 1986; Siculella and Palmer 1988; Fauron *et al.* 1989; Folkerts and Hanson 1989). White mustard and liverwort are exceptional for the relatively small size and unicircular form of their mitochondrial DNA (mtDNA) (Palmer and Herbon 1987; Oda *et al.* 1992a). At present, liverwort is the only plant species whose mtDNA has been proven, by direct electron-microscopic observation, to exist in the form hypothesized by its physical map (Oda *et al.* 1992a) and where complete nucleotide sequences are available for both chloroplast and mitochondrial DNA (Ohyaama *et al.* 1986; Oda *et al.* 1992b).

It has been reported that the mtDNA of *Brassica campestris*, for which one of the first ever physical maps of plant mtDNA was determined, contains a single pair of direct repeats, and that recombination between these repeated elements results in the division of the original circular molecule (218 kb) into two smaller molecules (135 kb and 83 kb). Consequently, the mitochondrial genome of *B. campestris* is thought to consist of three molecular species, an arrangement which has been termed a tripartite structure (Palmer and Shields 1984). This genome model seems to apply sufficiently well to much more complex genomes with a large number of repeated sequences active in homologous recombination. Subgenomic molecules generated by homologous recombination contain a certain region of the master circle (or master chromosome; Lonsdale *et al.* 1984). The master circle, therefore, represents an assemblage of the DNA molecules which are no longer able to generate further subgenomic molecules.

It is still unknown, however, whether the master circle in multipartite genomes really exists as a single DNA molecule or not, because no large circular DNA molecule corresponding to the master circle has ever been directly detected in the mitochondria of higher plants, by either electron microscopy or any other method (Bendich 1985; Bendich and Smith 1990). Rapid rearrangement of the mitochondrial genome, induced by prolonged cell culture or somatic cell fusion, suggests that the plant mitochondrial genome is in a dynamic equilibrium of homologous recombination (Lonsdale *et al.* 1988).

The molecular mechanism of mitochondrial gene expression is well understood in animals and fungi. In these organisms, genes are generally transcribed in a polycistronic fashion from a limited number of promoters, which have themselves been well characterized (reviewed in Shinkel and Tabak 1989). In vertebrates, there are only single transcription initiation sites for each strand of mitochondrial DNA (Attardi 1985), while several transcription initiation sites were found in the mitochondrial DNA of yeast, *Saccharomyces cerevisiae* (Christianson and Rabinowitz 1983; Biswas *et al.* 1987). At least two protein factors, RNA

polymerase and a transcription factor, were found to cooperatively initiate transcription in human, *Xenopus* and yeast mitochondria (reviewed in Schinkel and Tabak 1989).

Several molecular events in gene expression which are characteristic to plant mitochondria have been discovered. Some of the plant mitochondrial genes have multiple transcription initiation sites, and many of the transcription units often contain only a single gene. The location of the basal promoter sequence required for transcription initiation in plant mitochondria has been being confined to a short stretch of sequence motif, but its protein factors have not yet been identified (reviewed in Gray *et al.* 1992). Molecular mechanisms of other phenomena in plant mitochondria, i.e. RNA editing, *trans*-splicing, and tRNA incorporation from the cytoplasm, remain intriguing but entirely uncertain.

The major objective of this study is to provide fundamental information on the genetic system of rice mitochondria, rice being a crop of obvious importance, and the author focuses on two of its central aspects: genome structure and gene expression. In this study, a physical map of the rice mitochondrial genome was constructed (Chapter 1), and the structure and transcription of the three copies of the rRNA genes (*rrn18* and *rrn5*) were analyzed (Chapter 2), which provided a clue towards the elucidation of the so far ambiguous evolutionary pathway of rice (Chapter 3). These results are expected to contribute towards an understanding of the genetic system of rice mitochondria.

Physical Structure of Rice Mitochondrial Genome

Introduction

The expanded size and structural complexity of plant mitochondrial DNA (mtDNA) present a striking contrast to the concision of its animal counterparts. Despite these obstacles, physical maps of mtDNA were constructed in various plant species. Mitochondrial genomes from many plant species are constituted of several molecular species, which is generally regarded as a result of homologous recombination (Lonsdale *et al.* 1988). Several conspicuous reiterated sequences have been found in mtDNA of higher plants and are speculated to be involved in homologous recombination which produces a set of recombinant sequences. In plant mitochondrial genomes, there are also smaller repeated sequences which probably play a significant role in creation of much larger repeats and in rearrangement of plant mitochondrial genome (André *et al.* 1992).

Representing such a multipartite genome requires introducing a hypothetical molecule, 'master circle' or 'master chromosome', which includes overall sequence contexts divided into subgenomic molecules (Lonsdale *et al.* 1984). This model provides exceedingly rational explanation on both the molecular heterogeneity and the genetic integrity of the higher plant mitochondrial genome. The most serious flaw to this model is lack of evidence that a master circle does exist *in vivo* as a molecular entity and does contribute to maintenance and expression of the plant mitochondrial genome. The possibility of dispensing with the master circle as a molecular entity has been discussed by Lonsdale (1989), Folkerts and Hanson (1991), and in this chapter. Levy *et al.* (1991) and Narayanan *et al.* (1992) presented remarkable data which were obtained by an applied electrophoretic

technique and indicated existence of abundant subgenomic circular DNA molecules within mitochondria of maize suspension-cultured cell and rice. On the other hand, Bendich and Smith (1990) observed a predominant amount of linear DNA molecules from plant and yeast mitochondria by pulse-field gel electrophoresis. They observed no linear or circular molecules equivalent to a master circle. Thus, there must be more investigations to understand real molecular status of plant mitochondrial genome DNA.

The mtDNA from cultured cells of a CMS line (A-58CMS) of rice (*Oryza sativa* L.) was cloned and its physical map was constructed. The data presented in this chapter arise reappraisal of the conventional model which holds a single master circle as the general status of the mitochondrial genome of higher plants.

Materials and methods

Rice lines

The A-58CMS line has an *Indica*-type cytoplasm of the Chinsurah BoroII line on a *Japonica*-type nucleus of A-58 (a line suitable for relatively cold districts) and, therefore, exhibits alloplasmic CMS. Cultured cells in suspension derived from the A-58CMS and the Chinsurah BoroII lines were used as sources of mtDNA (Shikanai *et al.* 1987). Total DNA was prepared from leaves and shoots of the intact plants (A-58CMS).

DNA preparation and cloning

The isolation of mtDNA was performed as described by Shikanai *et al.* (1987) with a slight modification. The mitochondrial suspension was treated in lysis buffer with proteinase K at 65°C for 5 min prior to lysis.

To construct a mitochondrial genomic library, the isolated mtDNA was partially digested with *Mbo*I restriction endonuclease and ligated with the *Bam*HI digests of a cosmid pH79 for A-58CMS mtDNA and of a cosmid pWE15 for Chinsurah BoroII mtDNA, respectively, after the size fractionation of *Mbo*I digests by sucrose gradient ultracentrifugation (Maniatis *et al.* 1982). The ligates were

packaged using Gigapack Gold™ (an *in vitro* packaging kit of λ -DNA, Stratagene), then infected to *E. coli* HB101 or DH1 strain for A-58CMS mtDNA, and *E. coli* NM554 strain for Chinsurah BoroII mtDNA, respectively. Cosmid colonies obtained were transferred to microtitre plates for long-term storage and on nylon membrane (Biodyne A™, Pall) for colony hybridization.

The CTAB method (Lichtenstein and Draper 1985) was used to isolate total DNA from the leaves of intact A-58CMS plants and the shoots of A-58CMS grown in dark for 10 days at 25°C.

Hybridization analysis

DNA preparation was digested with restriction endonuclease and resolved by electrophoresis in 0.6% agarose gel. The DNA fragments fractionated in size were transferred to a nylon membrane (Biodyne A™ or Biodyne B™, Pall) by conventional capillary method. Hybridization was done at 42°C in hybridization buffer (50% formamide, 5×Denhardt's solution, 6×SSC, 0.5% SDS, and 0.2 mg/ml of sonicated- and denatured-calf thymus DNA) with a ³²P-labelled probe prepared by random priming reaction (a kit for the reaction was purchased from Boehringer Mannheim). When hybridization completed, the membranes hybridized with internally labelled probes were washed successively in 2×SSC containing 0.1% SDS at room temperature, in 1×SSC and 0.1% SDS at room temperature, and in 0.5×SSC containing 0.1% SDS at 65°C. Autoradiography was usually performed at room temperature with an X-ray film (Konica) but at -70°C with an intensifying screen when hybridization signals were faint.

Results

The restriction profiles of cosmid clones

Although it has been reported that tissue culture causes structural alteration of mitochondrial genome in plants (McNay *et al.* 1984; Rode *et al.* 1987; Chowdhury *et al.* 1988; Dörfel *et al.* 1989; Shirzadegan *et al.* 1989), we used the cell line in suspension culture as a source of mtDNA for two reasons; 1) the seeds of

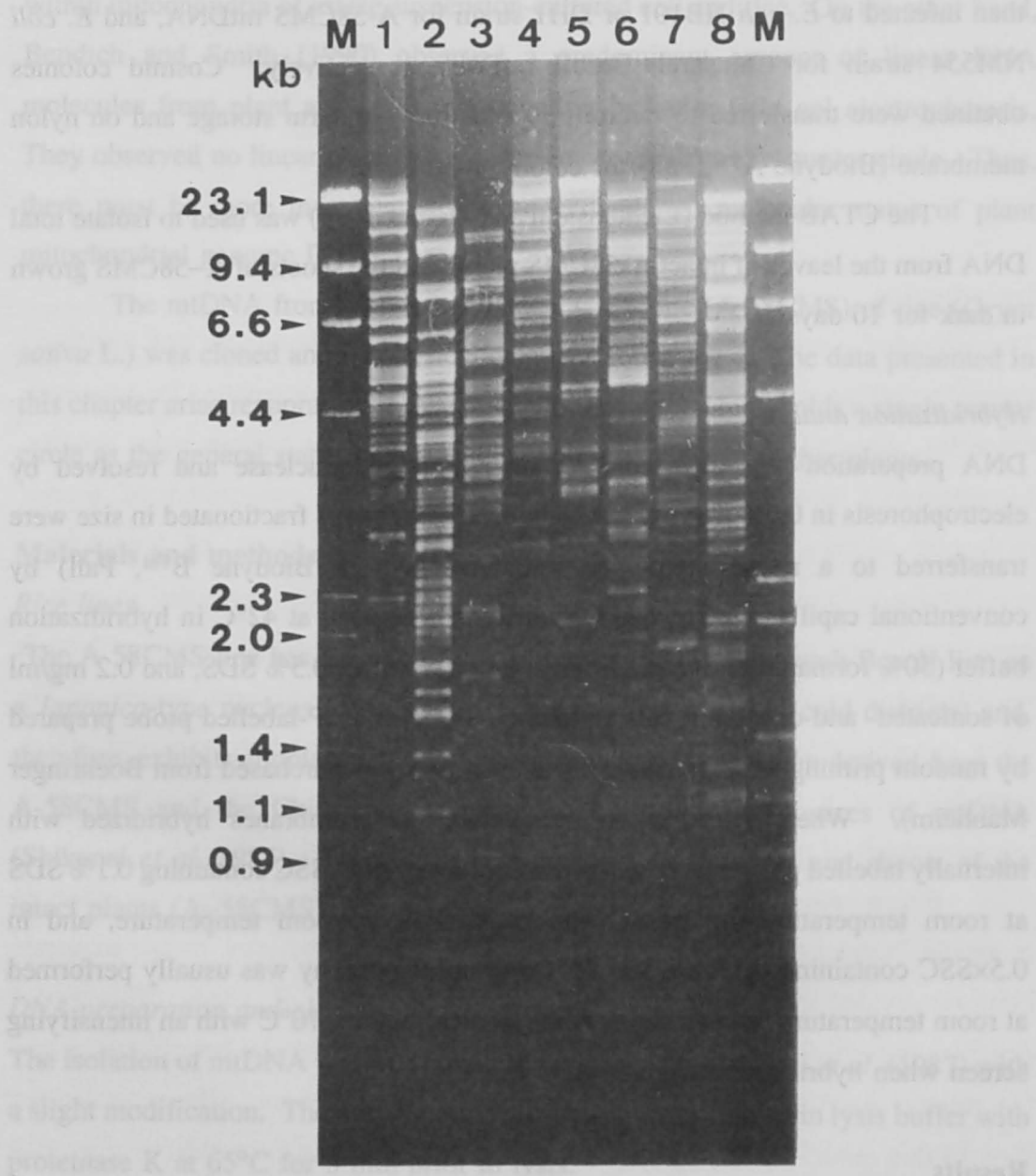


Fig. 1. Restriction profile of mtDNA prepared from the cultured A-58CMS. Mitochondrial DNA was digested with a respective restriction endonuclease: lane 1, *Bam*HI; lane 2, *Eco*RI; lane 3, *Hind*III; lane 4, *Kpn*I; lane 5, *Pst*I; lane 6, *Sac*I; lane 7, *Sma*I; and lane 8, *Xho*I. Lane M shows size markers (kb) of *Hind*III digested λ DNA and *Hae*III digested ϕ X174 DNA.

A-58CMS are of limited availability, 2) cultured cells must be retaining mitochondrial function necessary to maintain proliferation of cells. The isolated mtDNA from the cultured cells of A-58CMS line was digested with several restriction endonucleases. The mtDNA exhibited highly complicated restriction profiles (Fig. 1). Though multimolar restricted fragments make the size estimation difficult, the mitochondrial genome size was calculated at about 300 kb by summing up the sizes of the fragments.

The recombinant molecules containing a branching point

The *Xho*I, *Sma*I, and *Kpn*I restriction profiles of the obtained cosmid clones were compared with one another to establish overlaps between the individual clones. The cosmid library can be assumed to contain the whole mitochondrial genomic DNA because this library consists of all the major DNA fragments generated by digestion with the *Xho*I, *Sma*I, and *Kpn*I restriction endonucleases. The probes used to locate genes by Southern hybridization are summarized in Table 1. On the rice mtDNA, there are two distinct regions which are highly homologous to the 5' portion of the liverwort *nad2* gene, as well as a single region which shares homology with the 3' portion of the gene. The latter was found adjacent to the former, implying that there is a pair of repeated sequences homologous to the 5' portion of the *nad2* gene, or that the 5' portion of the *nad2* gene on the rice mitochondrial genome is split.

A number of divergence points (branching points) due to homologous recombination were observed on the restriction map of rice mtDNA during its construction with the cosmid library as described by Coulthart *et al.* (1990). Inserts of cosmid clones having a branching point can be attributed to either *in vivo* DNA recombination between homologous sequences, or cloning artifacts. However, it was possible to discriminate the latter by the Southern hybridization of a DNA fragment of a cosmid clone having a branching point to the total mtDNA digests or to the digested DNA from other cosmid clones (Lonsdale *et al.* 1986). Cosmid clones with a branching point were isolated. DNA probe having a gene (*rrn18*) for

Table 1 A list of known genes used for Southern hybridization

Gene	Source	Reference
<i>atp6</i>	<i>Oenothera</i> 1.5 kb <i>NheI</i> fragment	Schuster and Brennicke (1987a)
<i>atpA</i>	pea 1.5 kb <i>EcoRI-HindIII</i> fragment	Morikami and Nakamura (1987)
<i>cob</i>	wheat 0.7 kb <i>HindIII-BamHI</i> fragment	Quétier, personal communication
<i>cox1</i>	<i>Oenothera</i> 0.6 kb <i>BamHI-EcoRI</i> fragment	Hiesel <i>et al.</i> (1987)
<i>cox2</i>	<i>Oenothera</i> 0.8 kb <i>HpaI-PstI</i> fragment	Hiesel and Brennicke (1983)
<i>cox3</i>	<i>Oenothera</i> 1.1 kb <i>EcoRI-PstI</i> fragment	Hiesel <i>et al.</i> (1987)
<i>nad1</i>	<i>Oenothera</i> 1.1 kb <i>BamHI-EcoRI</i> fragment	Brennicke, personal communication
<i>nad2</i>	liverwort	Nozato <i>et al.</i> (1993)
<i>nad4L</i>	liverwort	Yamato <i>et al.</i> (1993)
<i>nad7</i>	liverwort	Nozato <i>et al.</i> in preparation
<i>rps13</i>	<i>Oenothera</i> 1.0 kb <i>EcoRI</i> fragment	Brennicke, personal communication
<i>rrn18</i>	wheat 1.1 kb <i>AvaI</i> fragment	Quétier, personal communication
<i>rrn26</i>	pea 1.8 kb <i>EcoRI</i> fragment	Nakamura, personal communication

a ribosomal 18S RNA showed multiple hybridized bands to the mtDNA digests of eight different restriction endonucleases (Fig. 2). This indicates that the *rrn18* gene exists as a multicopied gene. This was confirmed by comparing the physical maps of the cosmid clones isolated by colony hybridization with the *rrn18* probe (Fig. 3). There are three species (1A, 2A, and 3A in Fig. 3) of flanking sequences on one end of the common approximately 2 kb region containing *rrn18* gene, and four species (1B, 2B, 3B, and 4B in Fig. 3) on the other end of the flanking region. In fact, 11 of 12 possible recombinant sequences were obtained as cosmid clones (Table 2), the remaining one sequence must then be present in the mitochondria. Further analyses of cosmid clones revealed the presence of another recombination site close to the 2 kb region (*rrn18* gene) resulting in an additional flanking sequence (sequences 4B and 2B in Fig. 3). All of the four possible sequences on this recombination site were obtained, and it was confirmed by hybridization analysis that these recombinant sequences result from homologous recombination (data not shown). Therefore, the 2 kb region (*rrn18* gene) has three kinds of flanking regions on each end, indicating three copies of *rrn18* genes in the rice mitochondrial genome although it is not clear whether all the *rrn18* genes are transcribed. A three-copied *rrn18* gene has been found in rye mitochondrial genome where detailed physical mapping showed the presence of an outer and an inner repeat (Coulthart *et al.* 1990) and also in wheat mitochondrial genome (Quétier, personal communication). The rice mitochondrial genome is likely to

Table 2 Cosmid clones containing various configurations of *rrn18* gene

	1B	2B	3B	4B
1A	18S17	HB55	18S24	18S13
2A	18S14	18S29	18S43	18S10
3A	*	18S5	18S25	Sm4.0-1

* A cosmid clone was not obtained (see also Fig. 3)

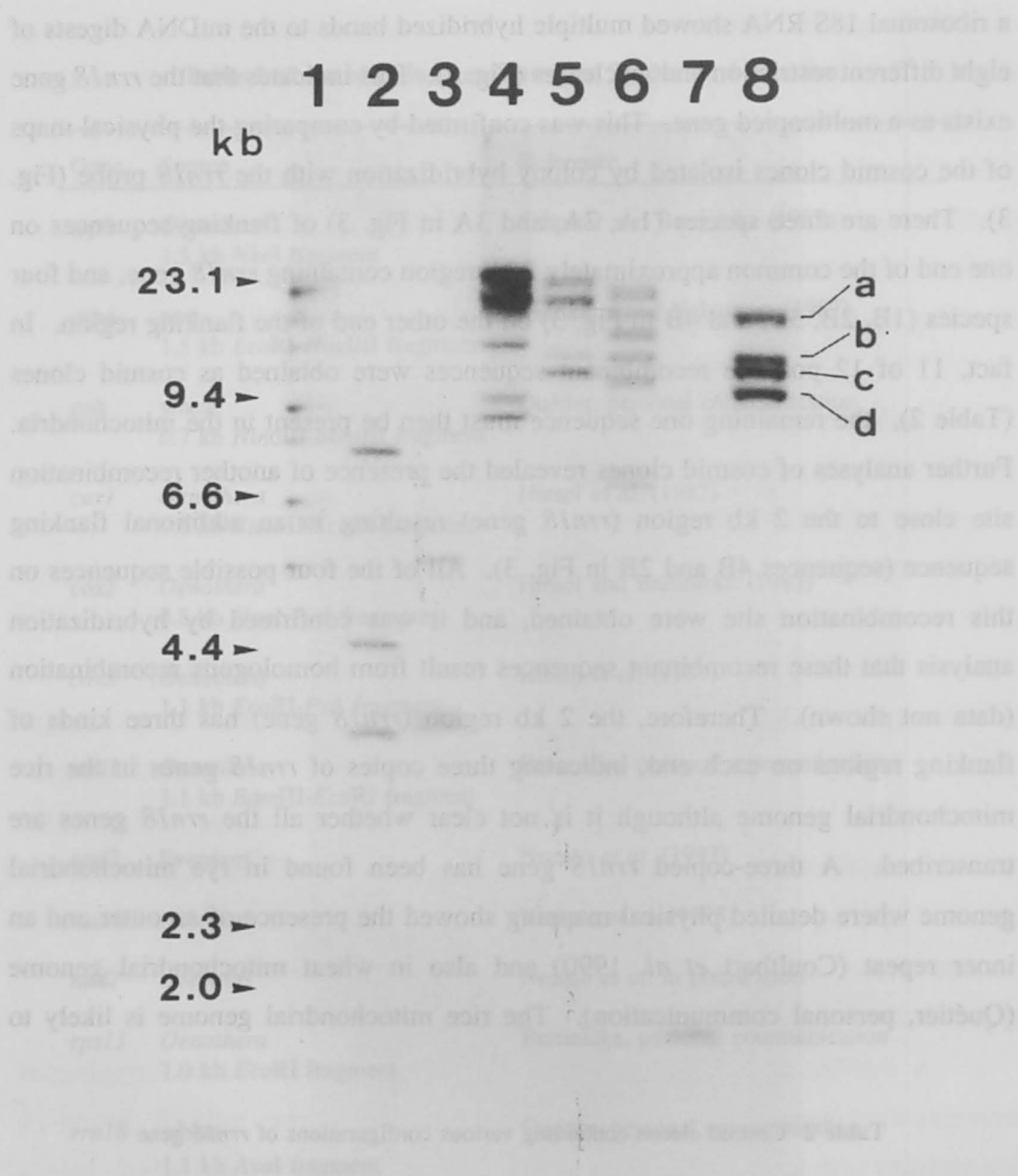


Fig. 2. Southern analysis of *rrn18* gene. DNA fragments shown in Fig. 1 were transferred to a nylon membrane and hybridized with a ^{32}P -labelled probe containing a portion of wheat mitochondrial *rrn18* sequence. The lanes are identical to those of Fig. 1. In lane 7, mtDNA digested with *Sma*I, showed a single band because this fragment is completely within a repeated sequence. All hybridization signals detected in *Kpn*I (lane 4), *Sma*I (lane 7), and *Xho*I (lane 8) can be attributed to the physical maps shown in Fig. 3. For example, the *Xho*I fragments (lane 8) with lower case letters (a, b, c, and d) correspond to the *Xho*I fragments shown in Fig. 3.

possess such kind of *rrn18* gene configuration indicating the close evolutionary relationship within the family, Gramineae. A three-copied repeat was found also in *Petunia* and nine possible recombinant sequences were detected (Folkerts and Hanson 1989).

As described earlier, a set of recombinationally-active repeats can have two branching points (X and Y in Fig. 4). For example, the branching point X has its partner Y with the repeat (R) indicating the existence of four kinds of recombinant sequences due to homologous recombination. In fact, two clones were obtained and the physical maps around the *atpA* gene are shown in Fig. 5A. The Southern hybridization with a probe of pea *atpA* gene to total mtDNA also showed two fragments having *atpA* genes corresponding to different clones (see the *Sma*I digested fragments in Fig. 5B). The branching point (corresponding to the

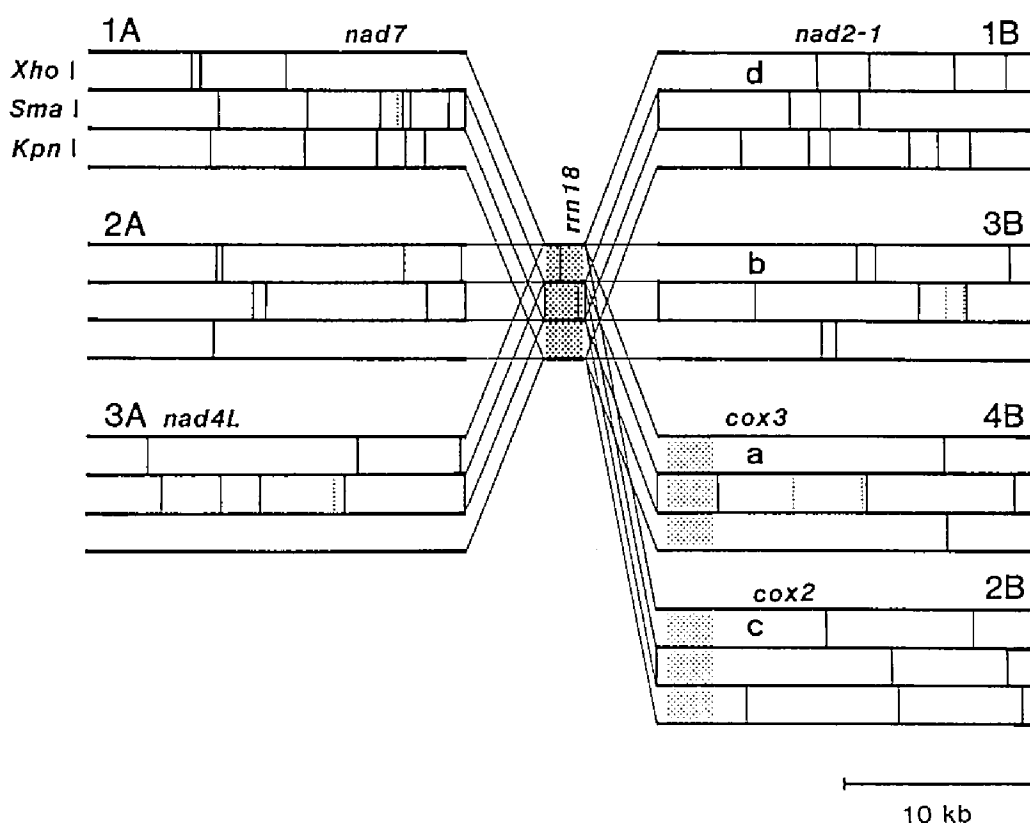


Fig. 3. Physical maps around *rrn18* gene. The densely-hatched region indicates a repeated sequence containing *rrn18* gene and the thinly-hatched region indicates another repeated sequence. The orders of fragments separated by dotted lines were not determined. The lower case letters (a, b, c, and d) correspond to the hybridization signals shown in Fig. 2.

branching point Y in Fig. 4) shown by a vertical line with a horizontal arrow was expected to be a product of homologous recombination due to repeated sequence near the *atpA* gene. This implies that another branching point (corresponding to the branching point X in Fig. 4) must exist somewhere in the direction shown with an arrow in Fig. 5A. To find out a partner of the branching point (so-called hidden branching point), the whole DNA of individual cosmid was labeled with ^{32}P -dCTP and hybridized to the total mtDNA digests of *Xho*I, *Sma*I, and *Kpn*I. Then any hidden branching points could be visualized as extra restriction fragments which are different from the restriction fragments of the cosmid DNA used as a probe (Fig. 6). However, we could not observe any signals indicating a hidden branching point in the region (approximately 60 kb long, the dashed line under linear map in Fig. 9) from the branching point near the *atpA* gene to the branching point due to the *rrn18* repeat. However, one hidden branching point was implied to exist near the *atp6* gene by performing the same experiments mentioned above (data not shown).

Comparison of mitochondrial gene organization from the cultured-cell and the intact plant

As described earlier, tissue culture can induce quantitative and qualitative alterations on plant mtDNA, and a nuclear genotype can also affect the structure of mitochondrial genome (Mackenzie *et al.* 1988; Escote-Carlson *et al.* 1990). We tried to find out whether the hidden branching point of the cultured A-58CMS

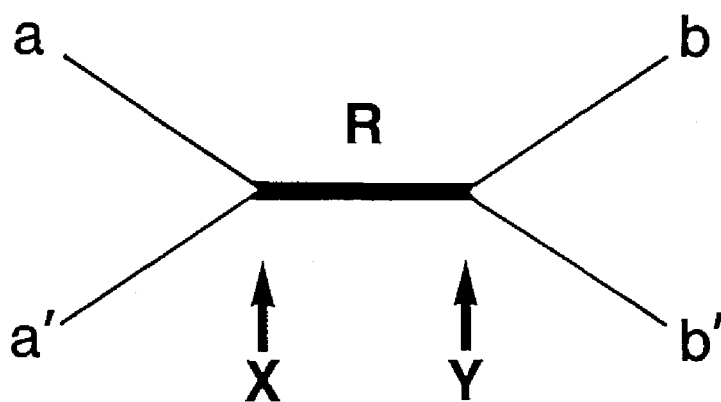


Fig. 4. Two branching points due to a pair of repeats. A diagram showing two branching points (X and Y) on the ends of the repeat (R) indicates the existence of four possible recombinant clones (a-R-b, a-R-b', a'-R-b, and a'-R-b').

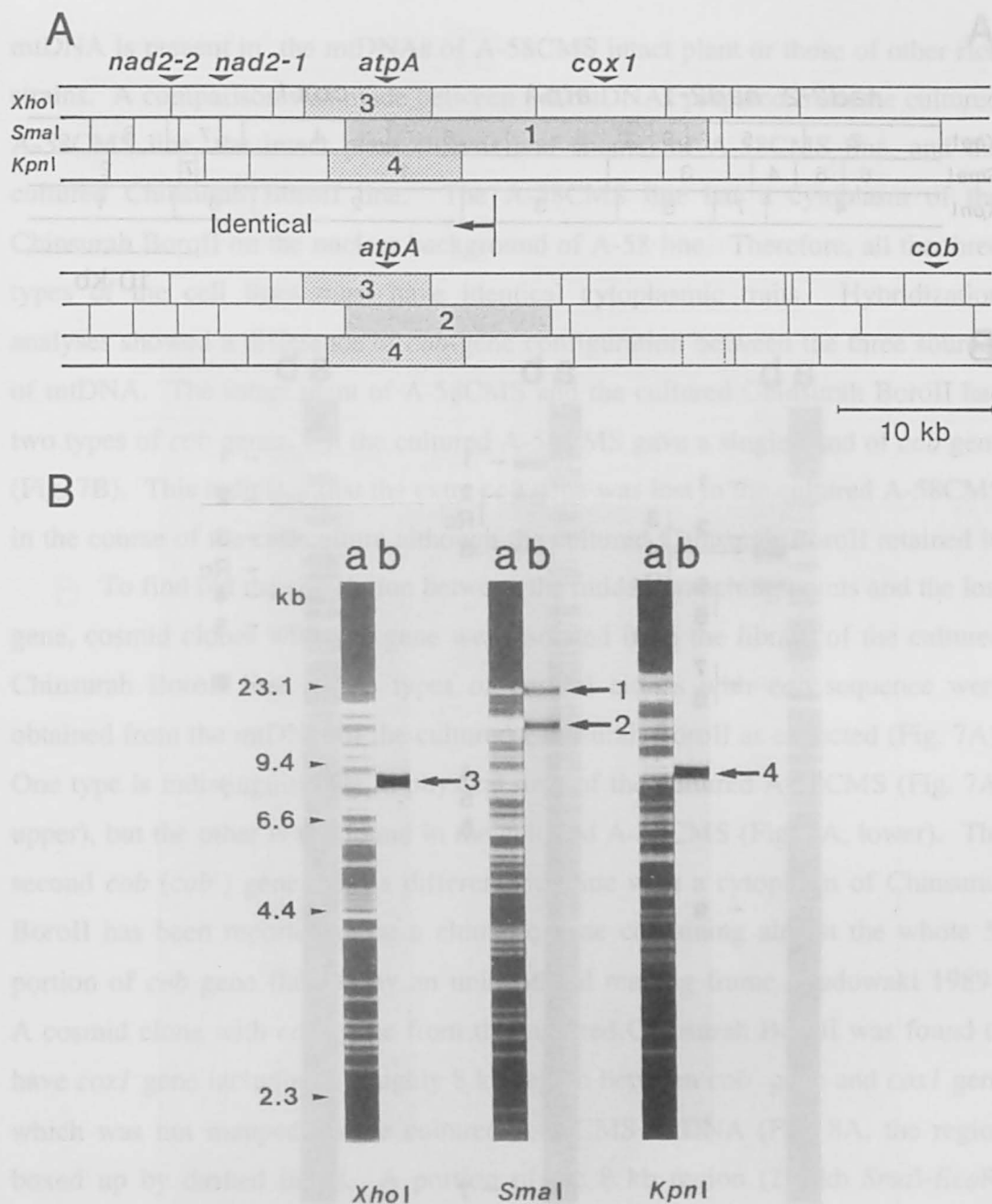
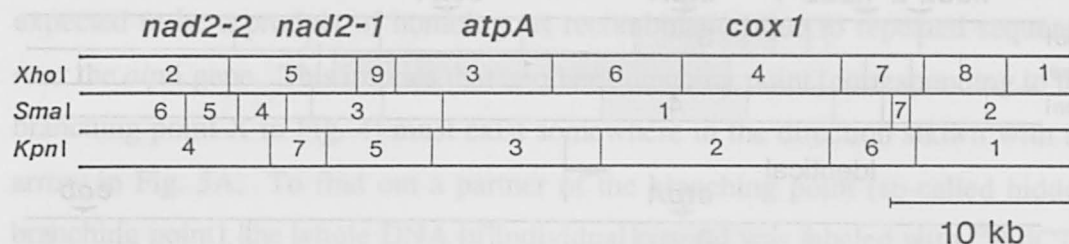


Fig. 5. A branching point adjacent to *atpA* gene.

A, Physical maps of two cosmid clones containing *atpA* gene. The approximate locations of *atpA* gene on the maps are shown. The dotted fragments (1, 2, 3, and 4) correspond to the hybridization signals indicated with arrows shown in **B**. A vertical line between the two maps is an approximate position of a branching point and the regions on the left of the vertical line are indistinguishable in physical map. **B**, Restriction profiles (lane a) and hybridization signals (lane b) of *atpA* gene. A probe containing a portion of pea *atpA* gene (Morikami and Nakamura 1987) was hybridized to the *XhoI*, *SmaI*, and *KpnI* digests of mtDNA.

A



B

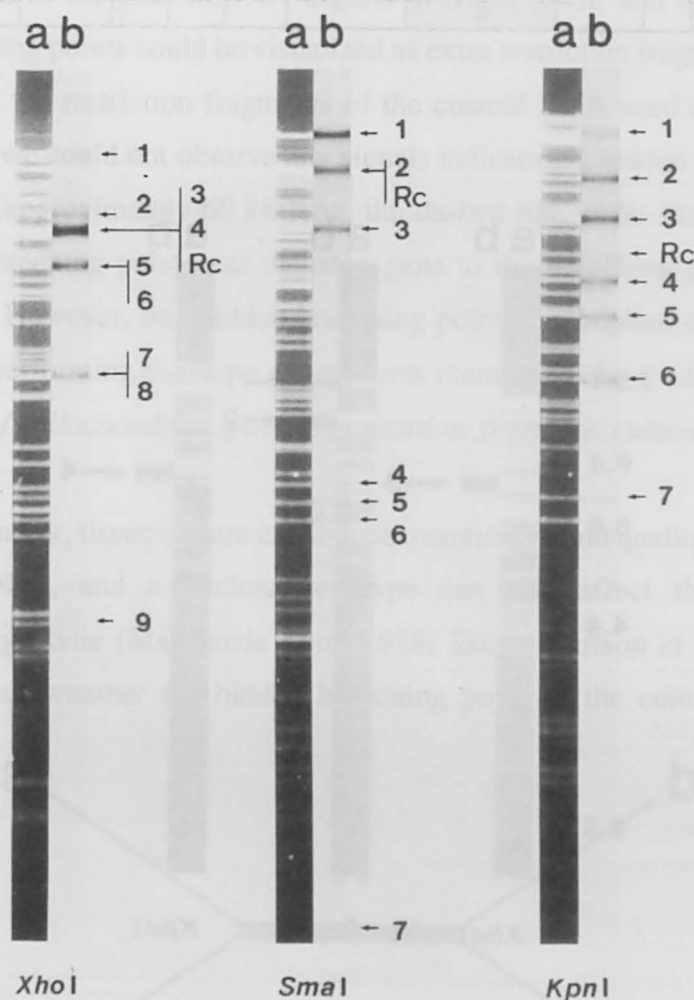


Fig. 6. Hybridization of a cosmid DNA (HB76, see Fig. 9) to the total mtDNA digests of the cultured A-58CMS. **A**, A physical map restricted with *XhoI*, *SmaI*, and *KpnI* of the mtDNA inserted in the cosmid clone (HB76) used as a probe. **B**, Mitochondrial DNA fragments digested with *XhoI*, *SmaI*, and *KpnI* (lane a) and hybridization signals (lane b) with a probe (32 P-labeled HB76 DNA). Numbers correspond to fragments shown in A. Rc indicates mtDNA fragment having the same branching point due to homologous recombination (see physical maps of clones in Fig. 5A).

mtDNA is present in the mtDNAs of A-58CMS intact plant or those of other rice strains. A comparison was made between the mtDNAs prepared from the cultured A-58CMS line, the intact plant (leaves and shoots) of A-58CMS line, and the cultured Chinsurah BoroII line. The A-58CMS line has a cytoplasm of the Chinsurah BoroII on the nuclear background of A-58 line. Therefore, all the three types of the cell lines must have identical cytoplasmic traits. Hybridization analyses showed a difference of *cob*-gene configuration between the three sources of mtDNA. The intact plant of A-58CMS and the cultured Chinsurah BoroII had two types of *cob* genes, but the cultured A-58CMS gave a single band of *cob* gene (Fig. 7B). This indicates that the extra *cob* gene was lost in the cultured A-58CMS in the course of the cell culture although the cultured Chinsurah BoroII retained it.

To find out the correlation between the hidden branching points and the lost gene, cosmid clones with *cob* gene were isolated from the library of the cultured Chinsurah BoroII line. Two types of cosmid clones with *cob* sequence were obtained from the mtDNA of the cultured Chinsurah BoroII as expected (Fig. 7A). One type is indistinguishable in physical map of the cultured A-58CMS (Fig. 7A, upper), but the other is not found in the cultured A-58CMS (Fig. 7A, lower). The second *cob* (*cob'*) gene from a different rice line with a cytoplasm of Chinsurah BoroII has been reported to be a chimeric gene containing almost the whole 5' portion of *cob* gene flanked by an unidentified reading frame (Kadowaki 1989). A cosmid clone with *cob'* gene from the cultured Chinsurah BoroII was found to have *coxI* gene including a roughly 8 kb region between *cob'* gene and *coxI* gene which was not mapped on the cultured A-58CMS mtDNA (Fig. 8A, the region boxed up by dashed lines). A portion of the 8 kb region (2.3 kb *SmaI*-*EcoRI* fragment shown in Fig. 8A) was hybridized to the mtDNA digests of the cultured A-58CMS and of the cultured Chinsurah BoroII. This result showed that the 8 kb region was missing from the mtDNA of the cultured A-58CMS (Fig. 8B). Therefore, the point from which the 8 kb region is derived corresponded to the hidden branching point in the cultured A-58CMS mtDNA (arrow in Fig. 8A).

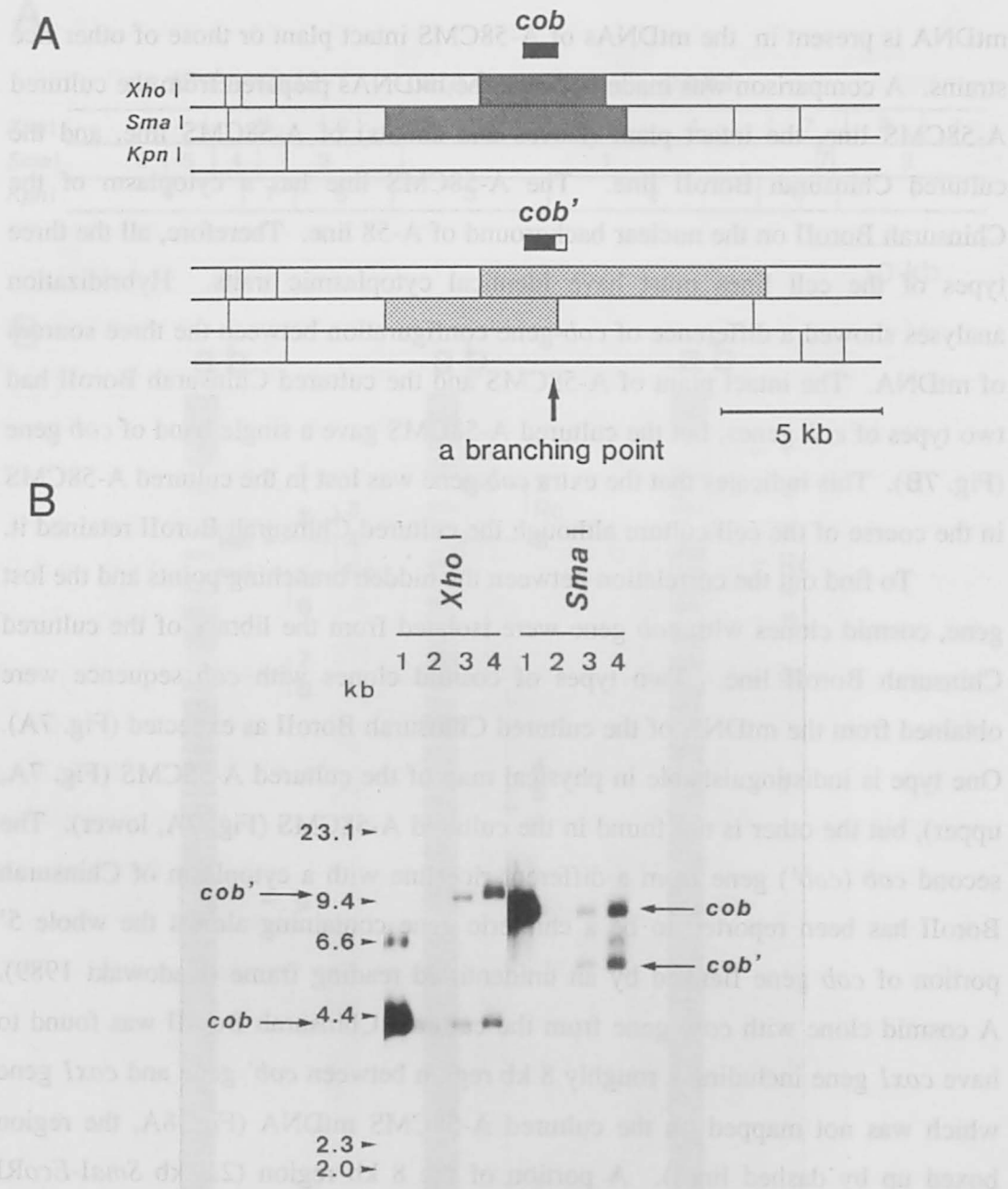


Fig. 7. A, Physical maps of two mtDNA regions of the cultured Chinsurah BoroII containing a sequence homologous to *cob* gene. The positions and length of *cob* and *cob'* genes (presented as a box above each physical map) were based on the data published by Kadowaki (1989) and Kaleikau *et al.* (1990). A closed box shows *cob* sequence and an open box shows an unidentified reading frame. Hatched restriction fragments correspond to hybridization signals shown in **B**. **B**, Southern hybridization of wheat *cob* gene to mtDNAs from the cultured A-58CMS (lane 1), total DNAs from mature leaves (lane 2) and shoot (lane 3) of A-58CMS plant, and mtDNA from the cultured Chinsurah BoroII (lane 4). The amount of DNA applied to lane 1 was much greater than those in lane 2, lane 3, and lane 4 to detect the *cob'* band seen in the other lanes. Signals at 6.6 kb may be non-specific.

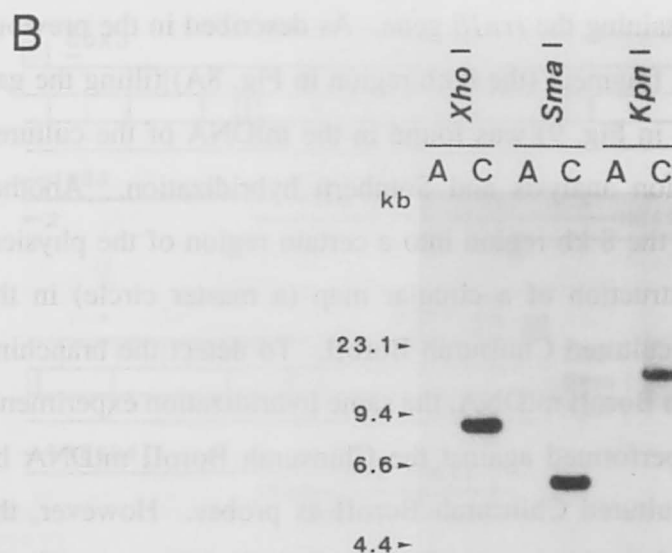
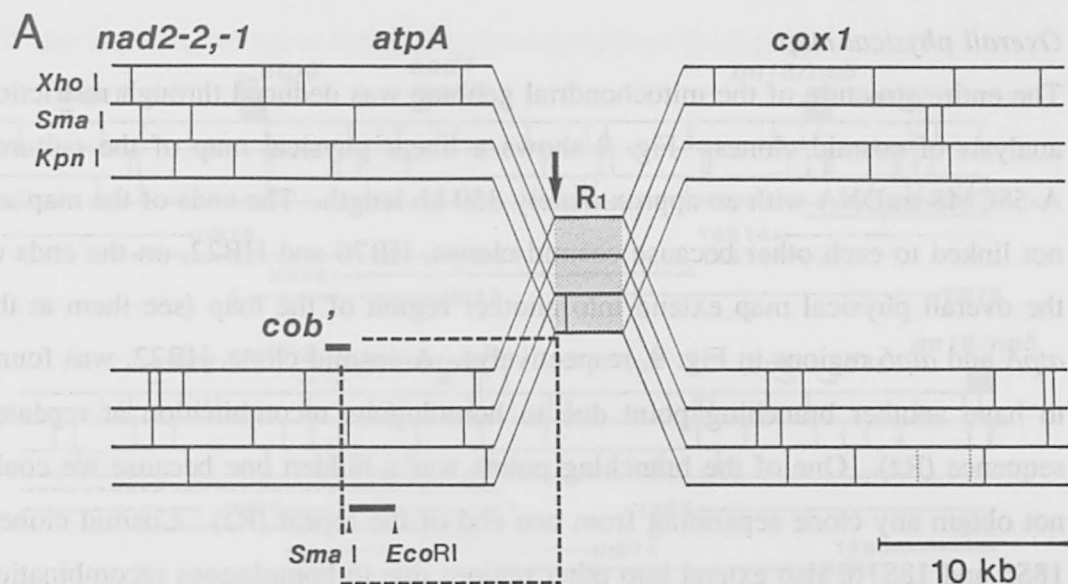


Fig. 8. **A**, Physical maps around a repeat (R1, a hatched region, its exact length is not determined) found in the cultured Chinsurah BoroiI. The 8 kb region is shown by a dotted box. **B**, Southern hybridization of a *cob'* specific fragment to mtDNAs from the cultured A-58CMS (lane A) and from the cultured Chinsurah BoroiI (lane C). An *Sma*-*Eco*RI 2.3 kb fragment of pOSB376 provided by Kadowaki (1989) was used as a probe (see the bar under the map). Arrow indicates a hidden branching point in mtDNA of the cultured A-58CMS.

Overall physical map

The entire structure of the mitochondrial genome was deduced through restriction analysis of cosmid clones. Fig. 9 shows a linear physical map of the cultured A-58CMS mtDNA with an approximately 350 kb length. The ends of the map are not linked to each other because cosmid clones, HB76 and HB22, on the ends of the overall physical map extend into another region of the map (see them at the *atpA* and *atp6* regions in Fig. 9, respectively). A cosmid clone, HB22, was found to have another branching point due to homologous recombination at repeated sequence (R2). One of the branching points was a hidden one because we could not obtain any clone separating from one end of the repeat (R2). Cosmid clones, 18S5 and 18S10, also extend into other regions due to homologous recombination at the repeated sequences containing the *rrn18* gene. As described in the previous section, a portion of the DNA fragment (the 8 kb region in Fig. 8A) filling the gap between the ends (R1 and R2 in Fig. 9) was found in the mtDNA of the cultured Chinsurah BoroII by restriction analysis and Southern hybridization. Another branching point derived from the 8 kb region into a certain region of the physical map is required for the construction of a circular map (a master circle) in the mitochondrial genome of the cultured Chinsurah BoroII. To detect the branching point in the cultured Chinsurah BoroII mtDNA, the same hybridization experiments as described in Fig. 6 were performed against the Chinsurah BoroII mtDNA by using cosmid DNAs of the cultured Chinsurah BoroII as probes. However, the remaining gap containing the branching point due to a hypothetical repeat (R2 in Fig. 9) was not found in the cultured Chinsurah BoroII mtDNA (data not shown). This indicates a possibility that the single master circle was absent even from the mitochondrial genome of the Chinsurah BoroII line.

Discussion

It is conventionally preferred to construct a plant mitochondrial genome into a single master circle. Most of mitochondrial genomes of higher plants so far investigated are reported to be circular possessing the whole genetic contents.

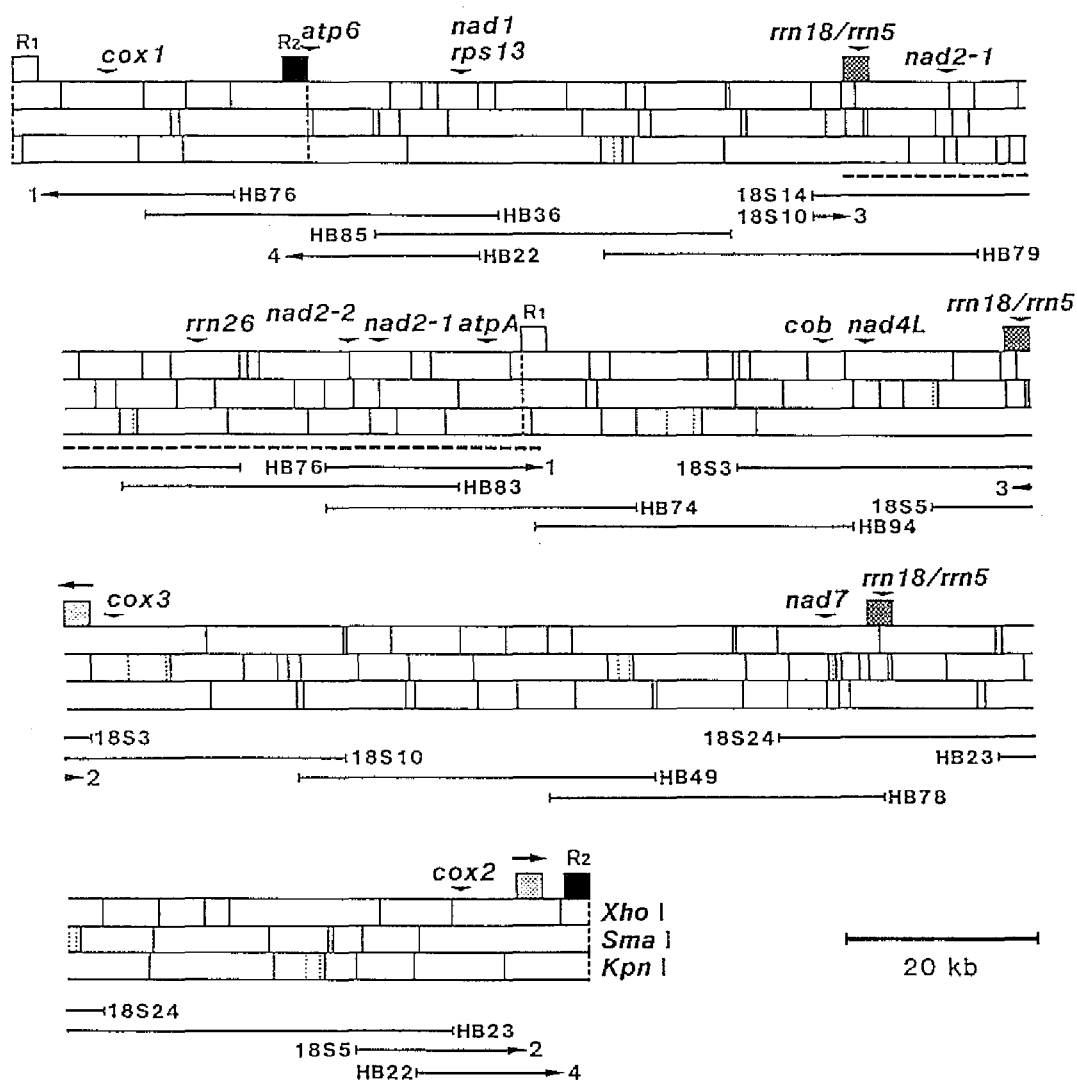


Fig. 9. The overall physical map of mtDNA of the cultured A-58CMS. Approximate locations of genes are shown by arrow heads with the genetic symbols. A minimal number of cosmid clones are shown by bars with their names under the physical map, each bar showing a region covered with a cosmid clone. The orders of fragments separated by dotted lines in the map are not determined. Due to homologous recombination, cosmid clones HB22, HB76, 18S5, and 18S10 have two separate regions on the map (bars with arrow and respective numbers). Dashed line under the map indicates the region searched for a hidden branching point for the repeat, R1 (see text). A variety of boxes over the map indicate locations of sets of repeated sequences. All sets of repeated sequences but one with arrows are oriented in same direction. Vertical dashed lines at the repeated sequences of R1 and R2 indicate arbitrary locations of a hidden branching point.

However, no direct evidence for the circularity of mtDNA has been presented. The results showed the missing of mtDNA fragment of the cultured A-58CMS line derived from the intact plant. A similar phenomenon was also reported in common bean at the restoration of fertility (Mackenzie and Chase 1990). Considering these results, construction of a single circular map with all cosmid clones requires an approximately 100 kb duplication that results in generation of a fourth *rrn18* gene (located in the thick arcs in Fig. 10A). On the other hand, two circular maps can be introduced to represent the whole genome to avoid the large duplication in the map (Fig. 10B). An elimination-duplication model for a single circle has been reported for two types of maize mitochondrial genome, the N genome (Small *et al.* 1989) and the V3 genome derived from the cultured cms-T cell (Fauron *et al.* 1990). The mitochondrial genome of the cultured A-58CMS is very similar to the

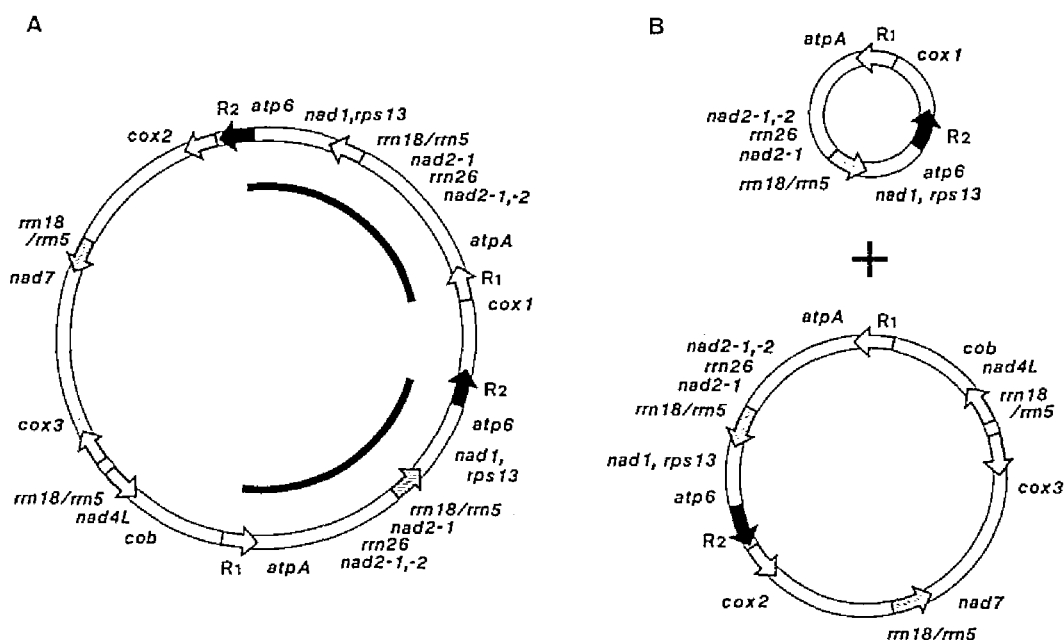


Fig. 10. The overall structure of mtDNA of the cultured A-58CMS. A variety of boxes with arrows indicate sets of repeated sequences. Arrows show relative orientation of the repeated sequences. **A**, A circular genetic map with large duplicated regions shown by thick arcs. **B**, An example of the dicircular model representing the whole mt genome.

V3 genome because the duplicated regions of the V3 (165 kb) and the cultured A-58CMS (100 kb) are just simply introduced into the respective genomes to make a circular map (thick arcs in the circle in Fig. 10A). It is unclear whether such duplications really exist in the genomes because the repeated sequences are too long (over 100 kb) to be cloned within a single cosmid. The validity of a physical map with such duplications is not definite. A DNA molecule corresponding to a master circle must be demonstrated by a physico-chemical method to construct the physical map of a single master circle. However, the duplications observed in the cultured A-58CMS, the N, and the V3 genomes may be introduced by a similar generation mechanism regardless of the lengths of duplications (approximately 100 kb in the cultured A-58CMS, 12 kb in the N, and 165 kb in the V3). Therefore, if an evidence of the 12 kb repeats in the N genome is demonstrated by the cloning of the DNA fragment containing the entire repeated sequences (12 kb), the duplication model for the other two genomes is possible.

Lonsdale *et al.* (1984) described the existence of 12 kb direct repeats and several other repeats on the N master circle, and Small *et al.* (1989) explained the occurrence of duplication of the 12 kb repeat containing the entire *atpA* sequence with a similar mechanism of duplication applied to the generation of the V3 genome. However, neither the cosmid clone having the full length of the 12 kb repeat and its both flanking regions in their master circle are obtained, nor the hybridization data indicating the presence of recombinant molecules are presented. It is both possible that cloning of the DNA segments is very difficult for an unknown reason or that the DNA segments are in fact absent from the DNA population of the N mitochondrial genome. Small *et al.* (1989) demonstrated the generation of the N master circle by combining two subgenomic circles derived from the master circle of an ancestral maize (RU) but the N master circle lacks a specific region for the RU genome.

Bendich and Smith (1990) are not in favor of the circularity of plant mtDNA and the existence of a master circle (a unit genome) as a genetic entity. Lonsdale *et al.* (1988) have mentioned the actual status of plant mtDNA to be difficult to

judge. As far as there is no appropriate technique for direct examination of DNA population and structure in higher plant mitochondria, a master circle can be a hypothetical molecule containing no more than a minimum and complete sequence complexity, i.e. net genetic content in plant mitochondria. Therefore, the mitochondrial genomes of the N maize, the regenerated cms-T maize, and the cultured A-58CMS line cannot be demonstrated as a single circle, rather the mitochondrial genomes must be shown as a linear genome with two ends of distinct repeats that caused elimination of the region between the repeats (Fig. 9). On the other hand, there can be a dicircular model where a genome comprises two subgenomic circles with a large duplication, if closed molecular feature or circularity of physical structure is definitely requested. Both of these representations seem to be insufficient compared with the conventional unicircular structure, but they are much more convincing to describe the structure of plant mitochondrial genome. The data described here cannot exclude the presence of a molecule corresponding to a master circle but does suggest the dispensability of a master circle as a molecular genetic entity.

How genetic integrity is maintained in apparent molecular fluidity of plant mitochondria is a matter of concern. Mitochondrial DNA in some organisms has been reported to be associated with the inner membrane, especially in the case of a slime mold, *Physarum polymorphalum* (Kuroiwa 1982). A complex is formed site-specifically with a membrane which may play a role in partitioning mtDNA in a dividing mitochondria (Kawano and Kuroiwa 1985). Fluorescence microscopic analysis revealed that mtDNA of *Allium cepa* exists as a nucleoid in a mitochondrion and segregates evenly into two daughter mitochondria during mitochondrial kinesis (Nishibayashi and Kuroiwa 1985). A mitochondrial nucleoid of yeast consists of nucleic acids and several proteins, one of which is basic and thought to bind to mtDNA (Miyakawa *et al.* 1987). Considering these facts, we can speculate how plant mitochondria maintain their genetic integrity. If a master circle is assumed, only its replication and distribution seem to be necessary and sufficient for the maintenance of genetic integrity although a strict mechanism for

protecting the master molecule from homologous recombination is required. If a protein complex to retain all the multipartite DNA molecules is assumed, a master circle molecule becomes non-prerequisite to the preservation of genetic information. If genes for proteins in the complex are nuclear-encoded and under control of nucleus, structural alteration of mitochondrial genome observed in cultured cells could be explained by substantial alteration of nuclear gene expression and organization induced by culturing procedure (Scowcroft and Larkin 1988). Furthermore, it is necessary for an extensively divided genome, a multipartite form, without a master molecule to have respective replication origins for amplification of the whole genome. Otherwise, a multipartite form without DNA replication origin can be extinguished, but no direct evidence of multiple replication origins has been obtained so far. On the other hand, active homologous recombination may be an alternative process for the proper DNA amplification to avoid the elimination of mtDNA segments. As discussed above, a mitochondrial genome is possible to exist in a multipartite form without a master circle.

Structure and Expression of the Three-copied *rrn18* Gene

Introduction

Three copies of the *rrn18* gene are present in the rice mitochondrial genome (Chapter 1), as in the wheat and the rye mitochondrial genomes (Coulthart *et al.* 1990, 1992). They exist in a set of molecular configurations which can be attributed to the effects of homologous recombination. Coulthart *et al.* (1990, 1992) demonstrated that the *rrn18/rrn5* repeats in wheat and rye consist of two outer repeats and an inner repeat which is included in the outer ones. Configuration of the rice *rrn18* repeats was apparently similar to those of wheat and rye, implying a close phylogenetic relationship (see Chapter 3).

Gene expression in plant mitochondria is not fully understood at present. Several transcription initiation sites have been determined for various genes of numerous plant species, and a consensus sequence for transcription initiation has been proposed (reviewed by Gray *et al.* 1992). A precise promoter sequence and other elements, if any, have yet to be elucidated, however. Recently, Hanic-Joyce and Gray (1991) have recently developed an accurate and efficient *in vitro* transcription system, prepared from wheat mitochondrial extract, which is expected to help clarify the factors required for transcription initiation.

The author was initially interested in examining the rice *rrn18/rrn5* repeats in order to gain an understanding of the process by which they were generated and the degree of the functional equivalence of the *rrn18* and *rrn5* genes on each repeated sequence. In the study described in this chapter, the nucleotide sequences of the rice *rrn18* and *rrn5* genes, including their flanking regions, and transcription initiation sites, were determined and characterized.

Materials and Methods

DNA sequencing and analysis

DNA fragments containing the *rrn18* gene were subcloned into pBlueScript (Stratagene), pUC118, or pUC119 from cosmid clones prepared previously (see Chapter 1). Sequencing was performed, in combination with the conventional nested-deletion and dideoxy methods, using Klenow fragment (Takara) and Sequenase (United States Biochemical). The sequences obtained were analyzed with the FLAT (Miyazawa 1991) and ODEN (Ina 1992) programs on a FACOM M-770 computer at the National Institute of Genetics, Mishima, Japan.

Mitochondrial RNA preparation

The mitochondria from a cell suspension cultured from the rice A-58CMS line, an alloplasmic, cytoplasmic male sterile line, were isolated by the method described by Shikanai *et al.* (1987) with the following minor modification: the DNase treatment prior to differential centrifugation was omitted. The purified mitochondria were lysed with 2% Sarkosyl in 50 mM Tris-HCl (pH 7.0) and 20 mM EDTA. To obtain the total mitochondrial nucleic acids, the lysate was extracted several times with a 1:1 mixture of phenol and chloroform, followed by precipitation with ethanol. The mitochondrial RNA (mtRNA) was separated from mtDNA by repeated LiCl precipitation (Ausubel *et al.* 1987).

S1 nuclease protection assay

The 0.6 kb *Sma*I-*Mlu*I fragment which covers the 5' flanking and coding regions of the *rrn18* gene (Fig. 1), was labeled at the 5' end of the *Mlu*I restriction site with [γ -³²P]ATP (Amersham) and T4 polynucleotide kinase (Takara). Ten μ g of purified mtRNA and 10⁴ cpm of the labeled probe were precipitated with ethanol. The pellet was redissolved in 30 μ l of hybridization buffer (80% formamide, 40 mM PIPES, 1 mM EDTA, 0.4 M NaCl). The solution was incubated at 75°C for 15 min, and then slowly cooled to 20°C over a period of 30 min, followed by incubation at the final temperature for 3 h. The nuclease protection assay was

performed in the presence of approximately 300 units of S1 nuclease, 200 mM NaCl, 30 mM sodium acetate (pH 4.6), and 10 mM ZnSO₄ to a total volume of 300 µl, at 20°C for 30 min. The reaction was stopped by adding 40 µl of 5 M ammonium acetate and 25 µl of 0.2 M EDTA, followed by phenol/chloroform extraction and isopropanol precipitation. The pelleted DNA was analyzed by denaturing polyacrylamide gel electrophoresis.

Primer extension

Two synthetic oligonucleotides with have appropriate sequences for determining 5' ends of the *rrn18* transcripts (Fig. 1) were labeled at their 5' ends as described above. Our protocol for primer extension was based on that of Sambrook *et al.* (1989), with several modifications. 10⁵ cpm of labeled primer and 10 µg of mtRNA were used per reaction, and the concentration of formamide in the hybridization buffer was set at 60% instead of 80%. The mtRNA hybridized with labeled primer was reverse transcribed in a reaction mixture (20 µl) containing 50 mM Tris-HCl (pH 8.0), 6 mM MgCl₂, 40 mM KCl, 0.2 mM dATP, 0.2 mM dGTP, 0.2 mM dCTP, 0.2 mM dTTP, 2.5 mM DTT, and 40 units of reverse transcriptase, at 42°C for 30 min; then 1 µl of the dNTP solution (mixture of dATP, dGTP, dCTP, and dTTP each at 1 mM) and 20 units of reverse transcriptase (AMV reverse transcriptase XL purchased from Life Sciences) were added to the mixture, followed by further incubation at 42°C for 30 min. The reaction product was precipitated with ethanol and redissolved in TE buffer.

Results

Comparison of the rrn18 and rrn5 gene sequences in rice and other plants

The mitochondrial *rrn18* gene of rice is much smaller than that of other plants shown in Fig. 1, because of a large deletion (approximately 250-300 bp) within variable region 4 (Fig. 2; nomenclature in accordance with Spencer *et al.* 1984). The nucleotide sequence of the *rrn18* gene is well conserved in several plant species; homology identities with rice are 92.5% for wheat, 91.6% for maize, 86.8%

[illegible]

rice CSU	GCACGTAGGCGGTGAATCGGGTTGAAAGTGAAAGTCGCCAAA--AAGTGGCGGAATGCTCTCGAAACCAATTCACTTGAGTGAG--ACAG	+709
wheat CSU	+709
maize	+702
soybean	+661
Oenothera	+664
liverwortCA.....GTC.....T.....T.....A.....A.....GC:T.....	+709
rice CSU	AGGAGAGTGGAAATTCGTGTGTAGGGGTGAAATCCGTAGATCTACGAAGGAACGCCAAAGCGAAGGCAGCTCTCTGGTCCCTA-CCGA	+798
wheat CSU	+798
maize	+791
soybean	+750
Oenothera	+753
liverwortA.....C:A:A.....G:AA:A.....T.....T.....TT:A:T.....	+799
rice CSU	CGTGGGGTGCAGAAAGCATGGGGAGCGAACAGGATTAGATACCCCTGGTAGTCCATGCCGTAAACGATGAGTGTCCGCCCTGGTCT----	+884
wheat CSU	+884
maize	+881
soybeanA.....G.....	+836
OenotheraG.....	+839
liverwortAAA.....A.....T.....TT.....	+885
rice CSU	-----AC-----	+886
wheat CSU	-----	+886
maize	ACGCTACGCT:	+893
soybean	-----	+838
Oenothera	-----	+841
liverwort	-----TGATATCGCACTCTTTGGGCTGACTTAAGCTCGGCTTAATGCTTAATTAAGTGAAGGAGTAGTGATCGATTGTT	+965
rice CSU	-----GCGGATCAGGGGCCAGCTAACGCGTGAAACACTCCGCTGGGGAGTACGGTGCAGAA	+943
wheat CSU	-----	+943
maize	-----	+950
soybean	-----T-----	+895
Oenothera	-----	+898
liverwort	TTCTTCAAGTTCACCAATCGTGAAATATGT:AT:.....TG:.....T:.....	+1055
rice CSU	GACCGAAACTCAAAGGAATTGACGGGGGCTGCACAAAGCGGTGGAGCATGTGGTTTAATTCGATACACCGGCAAAACCTTACCAGCCCT	+1033
wheat CSU	+1033
maize	+1040
soybean	+985
Oenothera	+987
liverwort	G:T:.....T:.....	+1145
rice CSU	TGCATATGAACAAACAACTGTCTTACGGGATGGTACTGACTTTCATACAGGTGCTGCATGGCTGTCTGCAGCTCGTGTCTGTGAGA	+1123
wheat CSU	+1123
maize	+1130
soybean	+1075
Oenothera	+1077
liverwortT:GTGTG:T:.....GA:AA:.....T:.....T:.....	+1235
rice CSU	TGTTTGGTCAAGTCTTATAACGAGCGAAACCCCTCGTPTTGTGTGGTGTGAGACATGCGCCTAAGGAGAAAT-----GC	+1193
wheat CSU	+1196
maize	+1203
soybean	+1165
Oenothera	+1167
liverwort	C:G:T:.....C:.....T:.....A:.....TTTGGTTC:ATCC:GACCCTGGAGACTGACGAA	+1325
rice CSU	CA-----CGAGTGACGTGCGCAGCGCTACTACTTGATTGAGTGCCA-----GCACGTAGCTGTGCTTTCAGCAAGAAATTCACCATTTGGGA	+1193
wheat CSU	+1277
maize	CGAGGAGCCGAGTGACGTGCCAGCGCTACTACTTGATTGAGTGCCA-----GCACGTAGCTGTGCTTTCAGCAAGAAATTCACCATTTGGGA	+1289
soybean	CGAGGAGCCGAGTGACGTGCCAGCGCTACTA-----ATTGAGCGCCA-----GCACGTAGCTGTGCTTTCAGTAAGAA-----GGGA	+1237
Oenothera	CGAGGAGCCGAGTGACGTGCCAGCGCTACTA-----ATTGAGTGCCAGCAAGCAGTGTGCTGTGCTAGTAAGAA-----GGGA	+1243
liverwort	GACTACGCCGTGAAATGGAGGATACCGACGAGCTAAGTTTTGCAACGCCGTGTGCTCATTCTGAAAGAAATATGACCATACAAAGT	+1415
rice CSU	GCCGCTGCCCT-----TCGAAGCACTTTCAGGTGTGAACCGAAGTGGCTTTGCC-----CAAGACCACGGAGAGCCTA-----	+1193
wheat CSU	+1345
maize	GCC-----TCGAAGCACTTTCAGGTGTGAACCGAAGTGGCTTTGCCGAACCTCAAGACCCACGGAGACCTA-----	+1354
soybean	GCCGCTGCCCTTTGGAATTCGAAGCACTTTCAGGTGTGGCTTTTGTGATTGCAAGTACGAGAGCAAGAAACCGATGCGGAAACGCGGC	+1327
Oenothera	GCCGCTGCCCTTTGCGAAACGTGTGGTATTGCAAGTACGAGCTTTCAGGTGTGGCTTTTCCCGAGCTGACA-----	+1303
liverwort	TTTCATACGGTATTCTCCGACGAACGAAGCTCTCAGAGCATTTG-----TACACTTCACT-----	+1471
rice CSU	-----CCTATAGTGACGTCAAAGTACCAGTGAGCATGGAGGTTTGGTTGAA	+1193
wheat CSU	-----	+1391
maize	-----CCTATAGTGACGTCAAAGTACCAGTGAGCATGGAGGTTTGGTTAGG	+1400
soybean	TTTCTTTGCGCTTGCTTGTGCTTTACTTAATAGAAAGAAAGGGCTTTTCGCTTTGTTAGTAAGTCCAGTTTGGCTTA--TCTT	+1415
Oenothera	-----GTTAGTAAGTCAAGTTTGGCCAC--TCTT	+1415
liverwort	-----	+1471
rice CSU	-----TGTTGTACGTAGTGGTAATAGTAC	+1217
wheat CSU	ATTGGTTACGACGACGTGAGTTTGGCGCGGAGGAAGACTCGGCATGAAGGCCAGAAA-----ATGG-----	+1477
maize	CTTGGTTACGACGACGTGAGTTTGGCGCGGAGGAAGACTCGGCATGAAGGCCAGGCCCGCCGCTG:.....A:.....	+1490
soybean	GCAGGTGACGACGACGTGAGTTTGGCGCGGAGGAAGACTCGGCATGAGTGAGGCCGCCCGCTG:.....GTT:.....	+1505
Oenothera	GCAGGTGATGACGACGTGAGTTTGGCGCGGAGGAAGACTCGGCATGAAGGCCAGGCCGCCCGCTG:.....GTT:.....	+1419
liverwort	-----GAGGAACCTGATCT-----CAA:GA:TG	+1500
rice CSU	GCGCCCC-----GCTCCGAACAAAGAAAAAGGTGCGTGCCTGACTACGAGGGACTGCCAGTGAGATCTGGAGGAAGTGGGGATG	+1300
wheat CSU	+1560
maize	+1572
soybeanC:G:CAAAAGC:.....CG:.....A:.....T:.....	+1595
OenotheraC:G:CAAAAGC:.....C:.....T:.....	+1508
liverwort	A:A:T:-----CCAT:.....T:.....A:AAA:.....T:.....	+1566

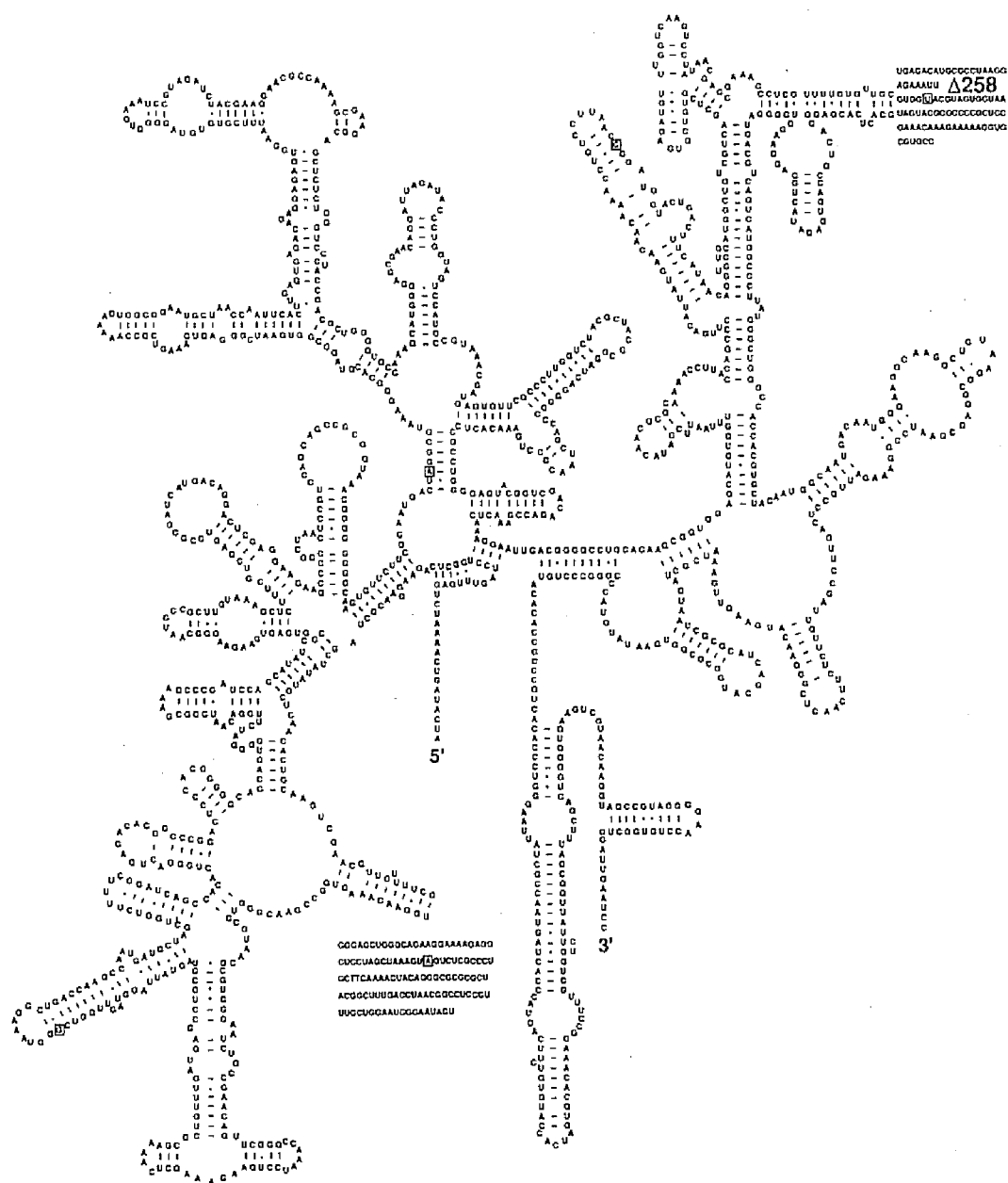


Fig. 2. Possible secondary structure of the rice mitochondrial 18S ribosomal RNA. Nucleotides which are substituted relative to the wheat counterpart are boxed. The deletion in rice is represented by a Δ symbol followed by the number of nucleotides deleted relative to the wheat *rrn18* gene.

for soybean, 88.8% for *Oenothera*, and 76.6% for liverwort. Only five base substitutions, discounting the large deletion, were observed when the rice sequence was compared to that of wheat, but these were not found to cause any alteration in the putative secondary structure of the 18S rRNA (Fig. 2). The nucleotide sequence of *rrn5* is also conserved among the species shown above. Though the 5' end of the rice *rrn5* gene could easily be predicted since this end region is identical to that of the wheat *rrn5* gene, the 5' and 3' ends of which have been experimentally determined (Spencer *et al.* 1981; Coulthart *et al.* 1992), the 3' end of the *rrn5* genes, and the downstream flanking region are not conserved between rice and wheat.

Structural analysis of the repeated regions

Nucleotide sequences of the thrice copied region of the rice mitochondrial genome, along with its flanking regions, were determined. No sequence divergence was observed among the three repeated sequences. A summarized map of the sequenced regions is illustrated in Fig. 3. There are two outer repeats and one inner repeat, 7213 bp and 5684 bp in length, respectively. The inner repeat was shared by all three of the repeats and was designated '18S/5S CSU' ('common sequence unit'), as in wheat and rye (Coulthart *et al.* 1992). The rice 18S/5S CSU is larger than that of wheat (4429 bp) and rye (2855 bp). For convenience, each boundary of the repeated sequences is termed according to its location relative to the *rrn18* gene, and flanking regions are assigned as shown in Fig. 3. The sequence coordinate +1 is assigned to the 5' end of the *rrn18* gene (Fig. 1). The 5' outer and the 5' inner boundaries, therefore, correspond to positions -1111 and -672, and the 3' inner and 3' outer boundaries to positions +5012 and +6102, respectively.

No previously identified gene, except for the *rrn18* and *rrn5* genes, was found within the rice 18S/5S CSU. As shown in Fig. 3, only an open reading frame coding for 187 amino acid residues (*orf187*) can be postulated to exist on the strand opposite from +3313 to +2750, and there is another open reading frame

(*orf152*) on the opposite strand crossing over the 3' inner boundary, from +5012 of the flanking regions 1B and 2B to +4662. The functionality of these two ORFs has not been studied. Several previously determined sequences (Fig. 3) were found by a computer-assisted homology search using DNA databases (Table 1).

The rice 18S/5S CSU begins with a sequence of 61 bp, homologs of which have been found in several other plant mitochondrial genomes (Fig. 4A). The location of the 61 bp sequence and its homologs in other plants is apparently not conserved. The *rrn18* genes of maize and soybean, for example, carry the homolog to the 61 bp sequence, but at different sites within the 5' flanking regions.

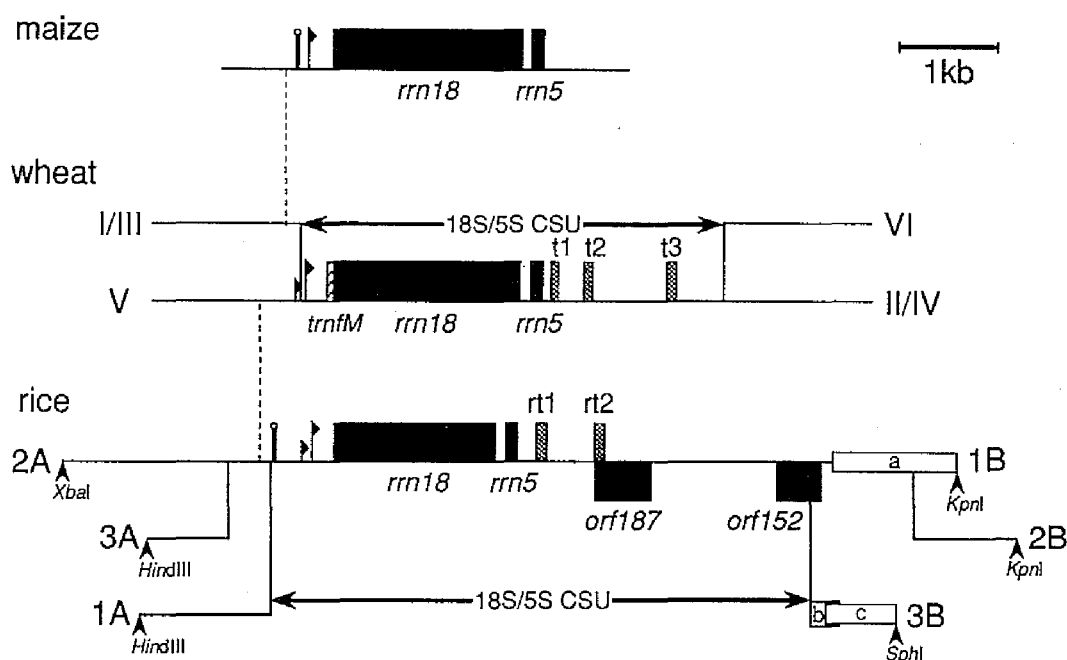


Fig. 3. Physical organization of the *rrn18* and *rrn5* genes in rice, wheat, and maize mitochondrial genomes. Genes and t-elements are represented by filled or hatched boxes above (rightward orientation) and below (leftward orientation) the horizontal lines. Open boxes are regions homologous to a, the rice variable copy number DNA, and b and c, the rice mitochondrial DNA sequences, pOST642 and pOSB945X, respectively, determined by Kadowaki *et al.* (1990). For simplicity, the other sequences given in Table 1 are not shown. Vertical lines topped with a triangle indicate transcription initiation sites; the longer ones are common to all three plant species, and the shorter ones are found only in rice and wheat. Vertical lines topped by a circle indicate conserved sequence elements (see Fig. 4). Each boundary of the repeated sequences is depicted by a vertical line. The vertical dashed lines indicate the upstream end of homology between the *rrn18* genes of the plant species shown. Restriction endonucleases which recognize the termini of the rice flanking regions are also given with arrowheads.

Table 1 Positions of sequences which are highly homologous to those in the GenBank database.

Homologous sequence	Position			Homology (%)	Reference
	From	To	Flank		
3' flanking sequence of the <i>rps13</i> gene from the wheat mitochondrial genome (from position 523 to 585)	-667	-616	CSU	79.7	Bonen (1987)
5' flanking sequence of the <i>cox1</i> gene from the rice mitochondrial genome (from position 204 to 262)	+1961	+2021	CSU	88.3	Kadowaki <i>et al.</i> (1989)
5' flanking sequence of the <i>trnP</i> (UGG) gene from the wheat mitochondrial genome (from position 459 to 709)	+2022	+2269	CSU	85.8	Joyce <i>et al.</i> (1988)
5' flanking sequence of the <i>trnC</i> from the tomato mitochondrial genome (from position 128 to 180)	-2515	-2466	2A	87.3	GenBank X13041
Variable copy number DNA from the rice nuclear genome (from position 1 to 1303)	+5234	+7167	2B	99.5	Kikuchi <i>et al.</i> (1987)
pOST642 (from position 30 to 338 : upper) and pOSB945X (from position 931 to 1671 : lower) cloned from rice mitochondrial DNA	+5003 +5170	+5310 +5911	 3B	99.5 99.9	Kadowaki <i>et al.</i> (1990)

Curiously, a secondary structure, shown in Fig. 4B, can be assumed for the rice 61 bp sequence and also for some of the other homologs. The sequences adjacent to the rice 61 bp sequence show a high degree of homology to flank V of the wheat 18S/5S CSU (Coulthart *et al.* 1992), while wheat lacks this structure.

Two tRNA-gene-like sequences were detected (rt_1 and rt_2) within the rice 18S/5S CSU (from positions +2123 to +2235, and +2739 to +2835, respectively). Similar structures have been detected downstream of the *rrn5* gene in wheat and rye and were designated as 't-elements' (Joyce *et al.* 1988; Coulthart *et al.* 1992). Wheat has three t-elements (t_1 , t_2 , and t_3) within its 18S/5S CSU, and rye while

A

source	position		reference
Os <i>rrn18</i> 5'	-668 -610	ACAAGCAACGGATTGAGCGcACTAGCGCGAAA-----GCCTTAG--cGCGCATCCGTTTCTTGCT	(61nt) 1
Zm <i>atp9</i> 3'	790 843	ACAAGCAACGGATTGAGCGc-----ATCCcTTGCGCCT-----GCCCATaCGTTTCTTGCT	(54nt) 2
Ta <i>tp613</i> 3'	370 423	ACAAGCAACGGATTGAGCGc-----ATCCcTTGCGCCT-----GCCCATaCGTTTCTTGCT	(54nt) 3
Ta <i>nad3</i> 5'	-228 -165	ACAAGCAACGGATTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(64nt) 4
Zm <i>rrn18</i> 5'	-401 -338	ACAAGCAACGGATTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(64nt) 5
Zm <i>OHF1</i>	-301 -253	ACAAGCAACGGATTGAGCA-CTA-----CGcTAG-----GCCCATCCGTTTCTTGCT	(48nt) 6
Gm <i>rrn18</i> 5'	-117 -58	ACAAGCAACGGATTGAGCA-CTA-----CGcTAG-----GCCCATCCGTTTCTTGCT	(60nt) 7
Ta <i>rrn26</i>	1813 1874	LgAAGCAAgGGA-TGAGC-----CGCAAAGCCGTTGCCCTTccgtaacCGCGCATaCGTTTCTTGCT	(62nt) 8
Zm <i>rrn26</i>	1828 1885	LgAAGCAAgGGA-TGAGC-----CGCAAAGCCGTTGCCCTTccgtaacCGCGCATaCGTTTCTTGCT	(58nt) 9
Zm <i>beta-R2</i>	4393 4329	TA-----AAGCCGTTGCC-----LgGCLcLATTCCGTTTCTTGCT	(34nt) 10
Ta <i>rrn26</i> 5'	-124 -187	gCAAGCAACGGATTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(64nt) 11
Zm <i>rrn26</i> 5'	-119 -162	gCAAGCAACGGATTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(64nt) 12
Os <i>rrn5</i> 3'	284 243	TTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(42nt) 1
Ta <i>rrn26</i> 5'	-207 -258	TTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(52nt) 8
Ta <i>trnP</i> 5'	-299 -350	TTGAGCA-CTAGCGCGAAAGCCGTTGCCCGTTA-----GCCCATCCGTTTCTTGCT	(52nt) 12
Gm <i>rrn18</i>	1346 1300	ACAAGCAAgGcgaaa-----GAAAGCC-----CGGTT-----CGCGCATCCGTTTCTTGCT	(47nt) 7
Ob <i>rrn18</i> 3'	2229 2174	AgAAGCAAgGGATgTAGCA-ACT-----LGGCGTataGCGTTA-----GCGCGATCCGTTTCTTGCT	(56nt) 13
Ll <i>rrn18</i>	1324 1304	ACAAGCAACGGATTGAGCAcACTAGCGCGAAAGCCGTTGCCCGTTA-----CGCGCATCCGTTTCTTGCT	(21nt) 14
Consensus		ACAAGCAACGGATTGAGCAcACTAGCGCGAAAGCCGTTGCCCGTTA-----CGCGCATCCGTTTCTTGCT (56nt)	

B

AA
G A
C-G
G-C
C-G
G-T
A-T
T-A
A-C-G
A C
C-G A
G-C
C-G
AG-C
G
TT-A
A-T
G-C
G-C
C-G
A-T
A-T
C T
G-C T
A-T
A-T
5'-A C-G C T-3'

Fig. 4. A, The rice 61 bp sequence and its homologs, all of which were found to be closely linked with genes in plant mitochondrial genomes. Abbreviations for the source organisms are as follows: Os, rice; Zm, maize; Ta, wheat; Gm, soybean; Ob, *Oenothera*; and Ll, lupine. The +1 position is assigned to the first nucleotide of each linked gene, though the sequence coordinate of the data in GenBank is given for the maize repeated sequence, β -R2, and relative locations to the linked genes are also given. Italic numbers indicate their complementary orientations. References 1, this work, 2, Bland *et al.* (1986), 3, Bonen (1987), 4, Gualberto *et al.* (1988), 5, Dale *et al.* (1985), 6, Hunt and Newton (1991), 7, Grabau (1985), 8, Falconet *et al.* (1988), 9, Dale *et al.* (1984), 10, Houchins *et al.* (1986), 11, Spencer *et al.* (1992), 12, Joyce *et al.* (1988), 13, GenBank X61277, 14, GenBank Z11512. B, Possible secondary structure of the rice 61 bp sequence. Normal base pairs are represented by a dash and GT pairs by a dot.

Ac									
===== <----->									
rt2	GTTCCTGG	CTGG--TTAGAGCAGAGGACTCGAAATCCTCTTTAG-----CTTTGCGCG							
rt1	::GAA::	A::GCAGT::G::GTG:::GT:::T::C:TTGCTGGGC::GG::C:T							
t1	:::::A	:::-:G:T-----AA							
t2	:::::	:::-CAGC::G::T:::GA:::C:-----CGG::CTT							
t3	:::::	:::-:-----GA:-----T:CA::A							
t4	:::::A	AA:::A:A-----G:::-----AAAAAGC::C:T							
t5	:::::	A:::-CT::GTG:::GT:::-----CGCTGG--:GG::C:T							

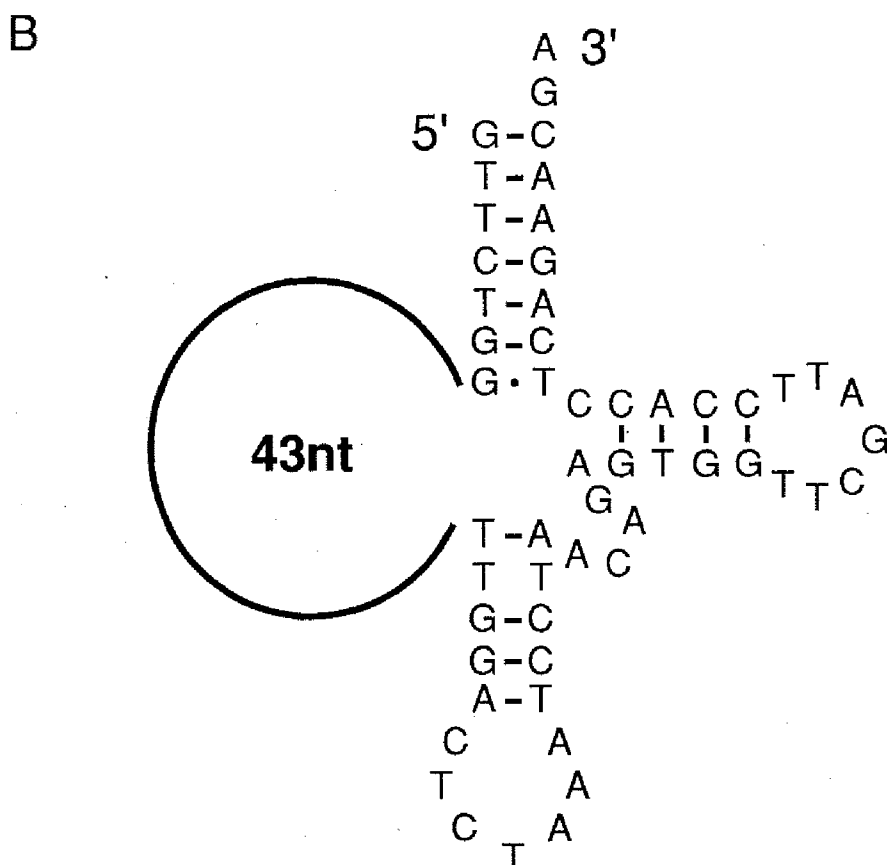


Fig. 5. **A**, Sequence alignments for tRNA-like structures. Nucleotides identical to those in rt_2 are represented by colons in rt_1 and wheat t-elements. Deletions are shown by dashes. Abbreviations: Ac, acceptor stem, An, anticodon stem, Ts, T ψ C stem, Tl, T ψ C loop, Vl, variable loop. **B**, Possible secondary structure of the rt_2 .

having only one in its 18S/5S CSU (t_1) has another within flank VI. Wheat has been reported to possess other t-elements upstream from the *trnP* genes (t_5 ; Joyce *et al.* 1988) and downstream from the *rrn26* gene (t_4 ; Spencer *et al.* 1992). As shown in Fig. 5, the primary and secondary structure of the t-element is moderately conserved and displays several characteristics typical of a tRNA gene, though some structural disorder can be seen in the stems due to base substitutions. The rt_1 resides within a region homologous to the wheat mitochondrial region containing the *trnP* gene (see Table 1), implying that rt_1 is the rice equivalent of the wheat t_5 . Although t_5 has intact stems (except for the D-stem) the acceptor stem cannot be organized in the rt_1 . The approximately 60 bp region upstream of the rt_2 shares homology with that of the wheat t_2 . The other intervening sequences (between the *rrn5* genes and the t_1/rt_1 , and between the t_1/rt_1 and the t_2/rt_2) show no homology between wheat and rice.

Transcription initiation site

For maize, the transcription initiation site for the *rrn18* gene is located at position -225 (the bold 'T' in Fig. 1; Mulligan *et al.* 1988b). The sequences flanking this site were conserved in rice and wheat. The conserved region contains the consensus sequence for transcription initiation in the maize mitochondria, i.e. [AT]CRTA[GT]A[AT]AAA (Gray *et al.* 1992; brackets indicate that nucleotide is possible at that position, and R signifies purine nucleotides), so that, based on the close phylogenetic relationship between rice and maize, transcription of the rice *rrn18* gene was, therefore, expected to begin at this position (position -222). To confirm this, the 5' terminus of the primary transcript was determined.

S1 nuclease analysis using a 5' labeled DNA probe complementary to the 0.6 kb region between the *Sma*I and *Mlu*I sites (Fig. 1) was performed, and three major signals observed. The strongest signal can be attributed to the 5' terminal of the mature 18S rRNA (Fig. 6). The other two signals were thought to correspond to the 5' ends of immature transcripts of the *rrn18* gene and were located at approximately positions -200 and -300.

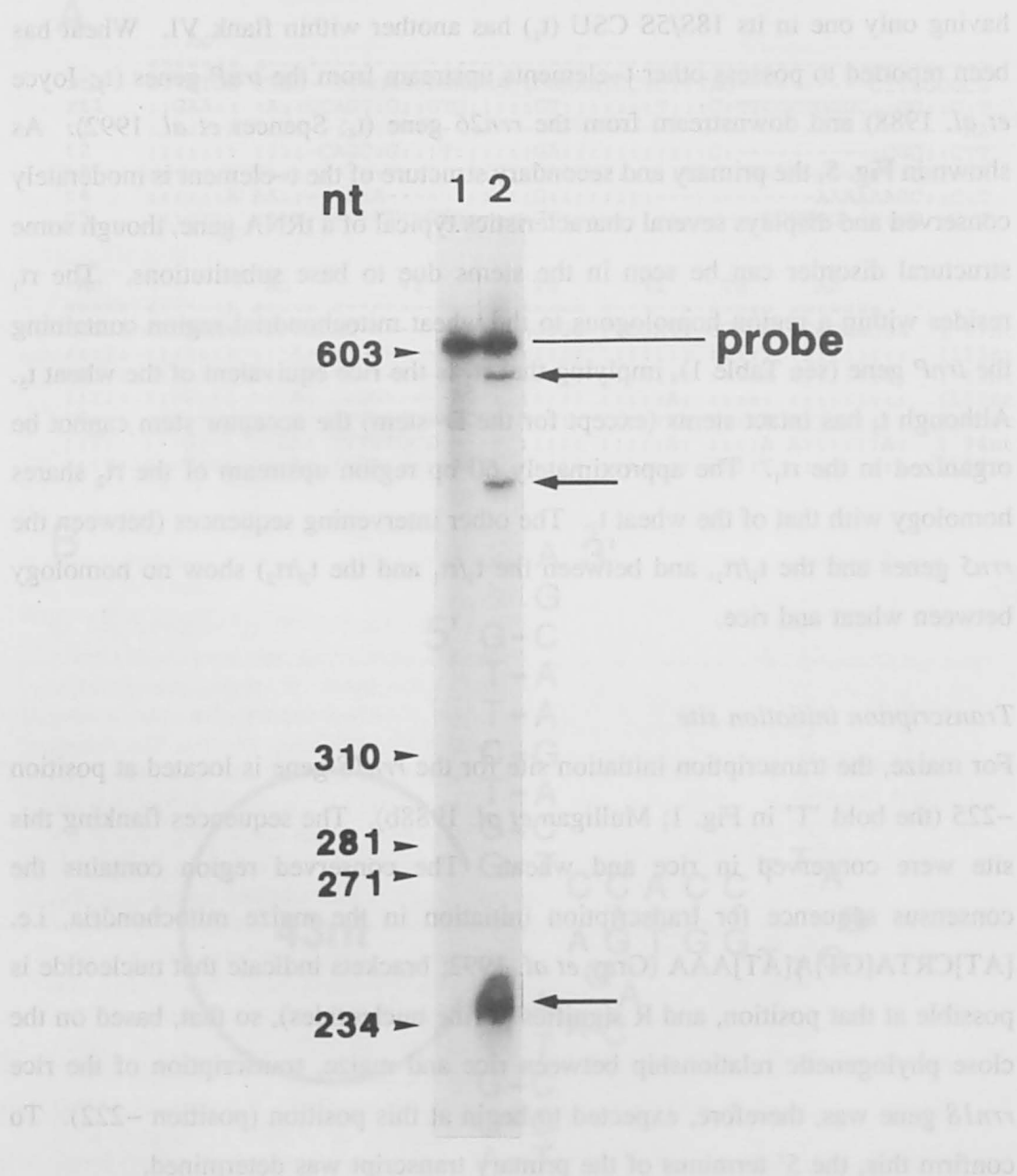


Fig. 6. S1 nuclease protection assay to determine the 5' ends of the *rrn18* transcript (arrows). The 617 bp *SmaI*-*MluI* fragment was asymmetrically labeled at the 5' end of the *MluI* site and used as a probe. The probe was treated with S1 nuclease in the presence (lane 2) or the absence (lane 1) of rice mtRNA.

The precise 5' terminus of each *rrn18* gene transcript was identified by primer extension (data not shown). The 5' terminus of the mature 18S rRNA was located at the identical site determined in wheat (Spencer *et al.* 1984). The 5' termini of the immature transcripts were confirmed to exist within similar sequence motifs ('CATAGAGAA' and 'CATAGCAAA'; 5' termini are shown in bold letters), which are homologous to the consensus sequence for transcription initiation in maize mitochondria. The base identities of the 5' termini are G and A, instead of T as in maize (Mulligan *et al.* 1988b), but their relative positions within the conserved sequence are identical (Fig. 1 and Table 2). This strongly suggests that the sequences, 'CATAGAGAA' and 'CATAGCAAA', are involved in transcription initiation in rice mitochondria.

Transcription initiation sites for several plant mitochondrial genes have been determined, though some of them require verification via analysis using a capping enzyme, which is able to discriminate primary (nascent) and secondary (processed) 5' transcript termini. Only transcription initiation sites whose surrounding nucleotide sequences are relatively conserved are listed in Table 2, but other transcription initiation sites without apparent sequence homology have also been reported (reviewed by Lonsdale 1989). The degree of sequence conservation among the sequences, especially 'CRTA', suggests that a critical role is played by the sequence motif in transcription initiation within plant mitochondria. Curiously, on the other hand, the precise sites where transcription starts are not conserved even within a single species.

Discussion

Structural features of the rice 18S/5S CSU and its flanking regions

Sequence analysis performed by the author demonstrated that all of the 18S/5S CSUs have identical primary structure in rice, probably due to the suppression of sequence divergence caused by homologous recombination.

Comparative analysis of the 18S/5S CSUs of rice, wheat, and rye, and the maize equivalent, revealed a high degree of conservation in primary structure. The

Table 2 Nucleotide sequences surrounding the transcription initiation sites of plant mitochondrial genes

Plant	Gene	Sequence ^a	Reference
Monocots			
rice	<i>rrn18</i>	CATAGAGAA	This study
	<i>rrn18</i>	CATAGCAA	This study
	<i>cob</i>	CATATAGAA	Kaleikau <i>et al.</i> (1992)
Wheat	<i>atpA/atp9</i>	CGTATAA AG	Covello & Gray (1991)
	<i>cox2</i>	CGTATAG TA	Covello & Gray (1991)
	<i>orf25</i>	CGTATAG TA	Covello & Gray (1991)
	<i>cox3</i>	CATAGAAT G	Covello & Gray (1991)
Maize	<i>rrn18</i>	CATAGATAA	Mulligan <i>et al.</i> (1988)
	<i>rrn26</i>	CGTATAAAA	Mulligan <i>et al.</i> (1988)
	<i>atpA</i>	CGTATTTAA	Mulligan <i>et al.</i> (1991)
	<i>atp6</i>	CATAGAGAA	Mulligan <i>et al.</i> (1991)
	plasmid S2	CATAAAGAT	Traynor & Levings (1986)
Dicots			
<i>Oenothera</i>	<i>cox1/cox3</i>	CGTA A GTGA	Hiesel <i>et al.</i> (1987)
	<i>atp6</i>	CATA A GTGA	Schuster & Brennicke (1987)
	<i>cox2</i>	CGTA T GTGA	Hiesel & Brennicke (1985)
Sugar beet	Mc.a ^b	CGTA A GTGA	Thomas (1992)
	Mc.c ^b	CATA A GTGA	Thomas (1992)
	Mc.d ^b	CATA A GTGA	Thomas (1992)
	<i>atpA</i> (CMS)	CGTA A GAGA	Thomas (1992)
	<i>atpA</i> (MF)	CATA A GAGA	Thomas (1992)
	<i>atp6</i> (CMS)	CGTA A GAGA	Thomas (1992)
	<i>atp6</i> (MF)	-ATA A GAGA	Thomas (1992)
Soybean	<i>atp9</i>	CGTAAGAGA	Brown <i>et al.</i> (1991)
	<i>scr b/c</i>	CATAAGAGA	Brown <i>et al.</i> (1991)
Consensus		CRTANNNNA	

Transcription initiation sites of the underlined genes were mapped by capping method.

^a Transcription initiation sites are indicated by bold letters.

^b Actively transcribed minicircular DNAs.

region from the possible transcription initiation site which is common to all four of the species, to the 3' end of *rrn5* gene, is highly conserved in nucleotide sequence, though there are some prominent insertion/deletion events, suggesting that the *rrn18* and *rrn5* genes are cotranscribed in rice, wheat, and rye as in maize (Maloney and Walbot, 1990). The most notable insertion/deletion event is the *trnfM* gene insertion in wheat and rye which exists upstream of the *rrn18* gene at a distance of a single nucleotide. If recombinational activity in plant mitochondria is assumed, it is a convincing hypothesis, as Coulthart *et al.* (1992) proposed, that a common ancestor of wheat and rye acquired the *trnfM* gene in one of the 18S/5S CSUs and then propagated it to the other CSUs with the assistance of homologous recombination. The approximately 40 bp sequence just upstream of the *rrn18* gene is well conserved in rice, maize, soybean, and *Oenothera*, while not in wheat and rye. This implies that this sequence may play a role in the processing of pre-18S rRNA transcripts. In wheat and rye, there is a tRNA gene (*trnfM*) instead of the approximately 40 bp sequence which is highly conserved among rice, maize, soybean, and *Oenothera*, and if the *trnfM* and *rrn18* genes are cotranscribed, processing of their tRNA results in the generation of the almost mature 5' end of the succeeding 18S rRNA.

The wheat t-elements are of interest because Hanic-Joyce *et al.* (1990) demonstrated that the wheat mitochondrial extract can properly process artificial RNA substrates containing the t-elements, t_1 , t_2 , and t_3 , *in vitro*, though no stable products of the t-elements could be detected *in vivo*. The rice t-elements, rt_1 and rt_2 , located downstream of the *rrn5* gene, are similar to the wheat t-elements. The spacer sequences separating the *rrn5* gene and the t-elements display no homology between rice and wheat, except for the short portion immediately upstream of the second t-elements, rt_2 and t_2 . It is not known whether the rRNA genes and the t-elements were brought together independently in rice and wheat, or share the origin; nor whether they were linked incidentally or towards some functional goal. If the t-elements have some function related to expression of the *rrn18/rrn5* genes, the t-elements following the *rrn18/rrn5* genes in rice and wheat would have

originated from a common sequence, and the spacer regions between the genes and the elements would then have been replaced with other sequences by homologous recombination between the t-elements and short sequences homologous to them.

Another interesting finding is a stem and loop structure which has been widely observed in the mitochondrial DNA of several plant species (Gualberto *et al.* 1988). In rice, this structure (61 bp) is located from position –670 to –610 of the *rrn18* gene, just inside (3 bp) the 18S/5S CSU. The relationship between the 61 bp structure and the 5' inner boundary is not clear at present. Compared with flank V of the wheat 18S/5S CSU, the structure is apparently a sequence inserted in the rice 18S/5S CSU, but the possibility of a deletion event in wheat can not be excluded. Wheat and rye lack the structure upstream of the *rrn18* gene, while maize and soybean do not. The wide distribution of this structure, and high degree of sequence similarity shown by it, imply an involvement in the mitochondrial genetic system, but its origin, function, and mechanism of propagation are entirely unknown.

The six flanking regions of the 18S/5S CSU (1-3A, 1-3B) were identified and analyzed in detail. Flank 1B was revealed to be nearly identical in nucleotide sequence (data not shown) to the variable copy number DNA of the rice nucleus (Kikuchi *et al.* 1987). The presence of both nuclear sequences in mitochondria (Schuster and Brennicke 1987b), and mitochondrial sequences in the nucleus (Fukuchi *et al.* 1991; Covello and Gray 1992) have been reported in plants, but there is not yet enough data to determine the transfer direction of the two rice sequences.

Transcription of the rrn18 genes in rice mitochondria

The two primary transcripts of the rice *rrn18* gene were found to have their 5' ends within the 18S/5S CSU, and both were located within the sequence motif, 'CATAGNRAA'. The first nucleotides (transcription initiation sites) of the two transcripts were mapped at the 7th nucleotide ('R', that is, 'G' or 'A', depicted by bold characters in the following motif sequences). The rice motif closest to the

rrn18 gene, 'CATAGAGAA', is identical to those of wheat and rye and has only one base substitution compared to that of maize, 'CATAGATAA'. The position of transcription initiation is conserved between rice and maize, but unknown for wheat and rye. The second rice motif, 'CATAGCAAA', which is located approximately 100 bp upstream from the first one, is also found in flank V of wheat, strongly suggesting that the wheat motif is also functional. The significance of this binary transcription initiation system is that it may, along with the multiplicity of the genes, give some advantage to rRNA supply. Wheat has multiple copies of *rrn26* gene, with two transcription initiation sites for each (Falconet *et al.* 1985; Spencer *et al.* 1992). The *rrn18* and *rrn26* genes are present in the maize mitochondrial genome as single copied genes (Lonsdale *et al.* 1984), and their transcription rates are the highest among all genes actively transcribed in maize mitochondria (Mulligan *et al.* 1991), thus showing simple correlation between copy number and transcription rate. In the case of rice, however, there is no information on transcription of the *rrn26* gene, and the relationship between substantial copy number and transcription rate of the rRNA genes is not clear in rice and wheat.

It is curious that, although the motif, 'CRTA', related to transcription initiation, is found within the upstream region of various genes in plant mitochondrial genomes, the exact transcription initiation site relative to this motif, and the identity of the first nucleotide, are not conserved even within individual species (Gray *et al.* 1992). What gives rise to this variability in transcription initiation is not known. The identical relative location of transcription initiation sites in the rice motifs, 'CATAGAGAA' and 'CATAGCAAA', suggests that a single molecular system is stringently involved in transcription initiation at these locations. This multiplicity of transcription initiation sites for the rice and wheat rRNA genes is different from that observed in the maize *atp9* and *cox3* genes, because of the apparent heterogeneity of transcription initiation sites in the latter (Mulligan *et al.* 1988a). Mulligan *et al.* (1991) presented intriguing data indicating that two transcription systems may be operating in maize mitochondria; stringent

transcription initiation signals for rRNA genes, *atpA*, and *atp6* genes, and permissive transcription initiation signals for *atp9* and *cox3* genes. More data on transcription initiation sites are required and the *in vitro* transcription system of plant mitochondria, as developed by Hanic-Joyce and Gray (1991), contributes greatly towards a more detailed comprehension of the transcription of plant mitochondrial genes.

The Evolution of Rice

Introduction

In spite of its importance to agriculture, the evolutionary relationship of rice, wheat, and maize has not yet been clarified. The shape of a phylogenetic tree involving these three plant species depends on how it is constructed. Clayton (1981), Enomoto *et al.* (1985), and Hamby and Zimmer (1988) presented a phylogenetic tree with rice as an outgroup on the basis of geographical distribution, restriction fragment patterns of chloroplast DNAs, and ribosomal RNA sequences, respectively. Wolfe *et al.* (1989) tentatively suggested wheat as an outgroup after comparing the nucleotide sequences of chloroplast genes, whereas Nugent and Palmer (1991) present maize as an outgroup. Evidence beyond that which rests on base replacement is therefore required to clarify the phylogenetic relationship of rice, wheat, and maize.

In the previous chapter, the author presented detailed structures of the rice 18S/5S CSU and its flanking regions, and results derived from comparative analysis of these structures provided evidence from which the shape of a phylogenetic tree for rice, wheat, and maize can be proposed.

Materials and Methods

Phylogenetic analysis was performed on a FACOM M-770 computer at the National Institute of Genetics, Mishima, Japan, with the ODEN program (Ina 1992). The ODEN program is capable of providing distance matrices via several methods, and can generate phylogenetic trees by the maximum parsimony (MP) method, the unweighted pair-group method with arithmetic mean (UPGMA), or the neighbour-joining (NJ) method. Nucleotide sequences were retrieved from the Genbank,

EMBL, or DDBJ databases.

Results

Comparison of the nucleotide sequences of the rrn18/rrn5 genes and their flanking regions.

The nucleotide sequences of the mitochondrial *rrn18/rrn5* genes and their flanking regions were determined for rice (Chapter 2), and, previously, wheat (Coulthart *et al.* 1992), rye (Coulthart *et al.* 1992), and maize (Dale *et al.* 1985). Ignoring some insertion/deletion events, such as the stem & loop structure in rice and maize and the *trnfM* gene in wheat and rye, the homology relationship of the sequences was arranged into a scheme, shown in Fig. 1.

Comparison of the nucleotide sequences of other genes

Phylogenetic trees for rice, wheat, and maize were constructed using the nucleotide sequences of genes for proteins and ribosomal RNAs, by the UPGMA and NJ method (Fig. 2). The evolutionary distances were calculated by a 6-parameter method at sites which are shared by all sequences analyzed. As shown in Fig. 2, shapes of the phylogenetic trees of genes varied so widely that it was impossible to determine the species phylogenetic tree from the results.

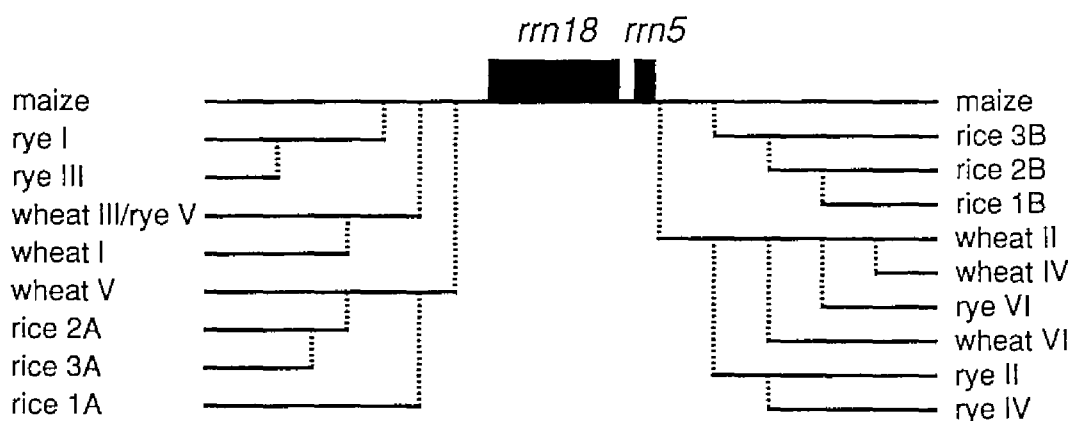


Fig. 1. Highly schematic diagram indicating the relationship of the flanking regions of the *rrn18* and *rrn5* genes. Vertical dashed lines indicate points of sequence divergence.

Discussion

If, based on nucleotide sequences, an identical shape is obtained for phylogenetic trees by both the UPGMA and NJ method, the resulting species tree is likely enough to be (Nei 1989). Nine of the genes analyzed showed consistency between the UPGMA and NJ method, but not among themselves; *rbcL* (chloroplast), *atpA* (chloroplast), *rpoA* (chloroplast), and *rbcS* (nucleus) presented the tree shape of Fig. 2B-1; *psaC* (chloroplast), *cox2* (mitochondrion), and *cox3* (mitochondrion), of Fig. 2B-2; and *petB* (chloroplast) and *nad3* (mitochondrion), of Fig. 2B-3. There is a

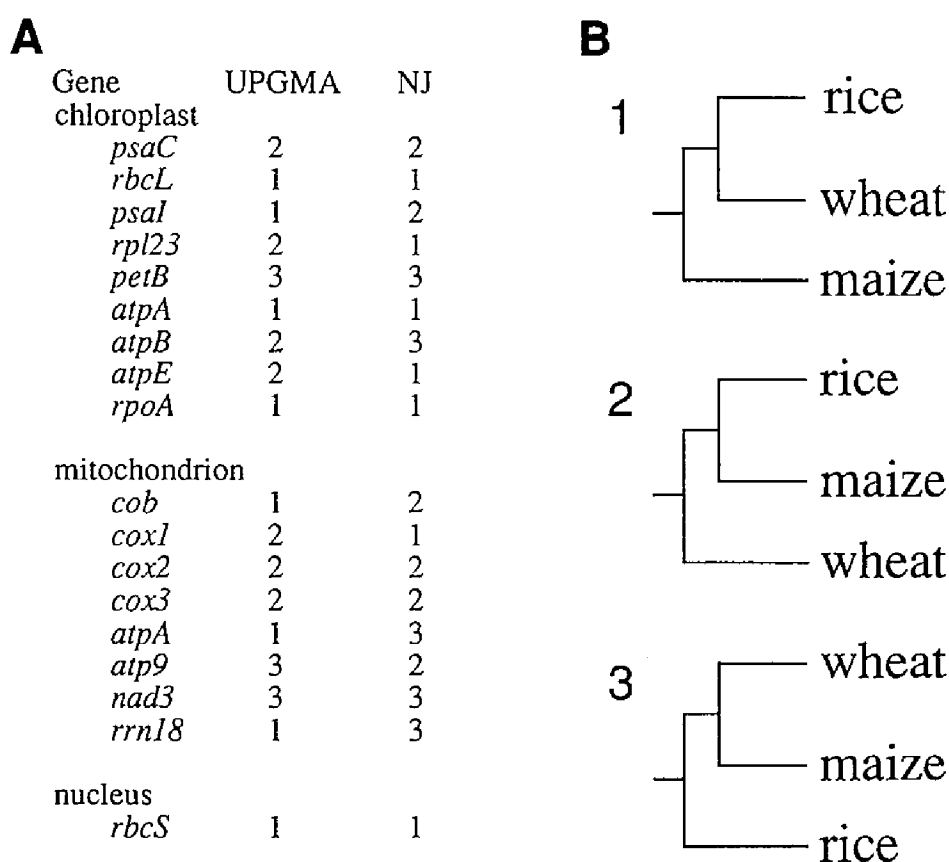


Fig. 2. Phylogenetic tree shapes based on differences in nucleotide sequences in the genes indicated. **A**, The genes listed were analyzed by both the unweighted pair-group method with arithmetic mean (UPGMA) and neighbour-joining (NJ) method. The shape of the phylogenetic tree for each gene is given by a number which corresponds to numbers given in **B**. Plant species used as an outgroup are tobacco for chloroplast and nuclear genes, and *Oenothera* for mitochondrial genes. Introns were excised and RNA editing ignored for the mitochondrial genes. Reference: see Table 1. **B**, Three possible patterns for the phylogenetic tree representing the evolutionary relationship of rice, wheat, and maize.

Table 1 GenBank Accession No. of the sequences analyzed in Fig. 2

Gene	rice	wheat	maize	tobacco	<i>Oenothera</i>
chloroplast					
<i>psaC</i>	X15901	M22132	X13159	Z00044	
<i>rbcL</i>	↑	X62117	J01423	↑	
<i>psaI</i>	↑	X62117	X61188	↑	
<i>rpl23</i>	↑	X12850	X07864	↑	
<i>petB</i>	↑	X54751	X05422	↑	
<i>atpA</i>	↑	M16842	X05255	↑	
<i>atpB</i>	↑	M16843	J01421	↑	
<i>atpE</i>	↑	M16843	J01421	↑	
<i>rpoA</i>	↑	X15595	X07810	↑	
mitochondrion					
<i>cob</i>	X53710	X02352	X00789		X07126
<i>coxI</i>	X15990	X56186	X02660		X05465
<i>cox2</i>	X01088	X01108	V00712		X00212
<i>cox3</i>	X17040	X52539	X12728		X04764
<i>atpA</i>	X51422	X16049	M16222		M24234
<i>atp9</i>	X16936	X15083	M18339		X15765
<i>nad3</i>	M57904	X14262	X14709		X52200
<i>rrn18</i>		see Fig. 1, Chapter 2			
nucleus					
<i>rbcS</i>	X07515	M37477	Y00322	X02353	

possible explanation for this inconsistency.

It is theoretically impossible to construct a phylogenetic tree from sequence data if genetic polymorphism occurred prior to divergence of each species, and if the period of species divergence is not long enough relative to the effective population size (Nei 1989). Therefore, if the first and the second divergences of rice, wheat, and maize occurred in a short period compared to their effective population size, a phylogenetic tree constructed for an individual gene would not reflect the true evolutionary pathway of the species.

One possible clue to the above problem lies in the copy number of the *rrn18/rrn5* genes in rice, wheat, and maize mitochondria. Rice and wheat possess three copies of the genes, while maize possesses only a single copy. The simplest interpretation of this evidence is that the evolutionary pathway of rice, wheat, and maize might be the one shown in Fig. 2B-1. There are, however, several problems

to be considered.

As shown in Fig. 1, the homology relationship of the flanking regions of the *rrn18/rrn5* genes in the rice, wheat, and rye mitochondrial genomes does not clearly reflect their evolutionary relationship. The origin of each flanking region must at least be identical between wheat and rye due to their high degree of relatedness (Coulthart *et al.* 1992), but Fig. 1 does not directly support this. A possible simple explanation for this matter exists, and is based on recombination between repeated sequences within plant mitochondrial DNA and the extinction of certain types of DNA molecules. On the assumptions 1) that three copies of the ancestral 18S/5S CSU could undergo frequent homologous recombination between one another, 2) that two sets of small repeats would locate closely upstream from the 18S/5S CSU and could induce homologous recombination as illustrated in Fig. 3, and 3) that the regions between the small repeats and the 18S/5S CSUs (the regions marked by asterisks in Fig. 3) would not be required for cell proliferation and that molecular species within the regions would have been extinguished during the course of evolution; then two of the immediate 5' flanking regions of the 18S/5S CSU can be replaced by the other 5' flanking region, as Small *et al.* (1989) initially proposed to explain the generation of the maize 12 kb repeats. The initial and final genomic structures of this process are depicted as circles corresponding to master chromosomes only for convenience and simplicity. The identical mechanism is also applicable to the mitochondrial genomes of wheat and rye, and also to the downstream flanking regions of the 18S/5S CSU. In partial support of this model, several analyses focused on the physical maps of plant mtDNAs have shown that the plant mitochondrial genome evolves rapidly (Palmer and Herbon, 1988), and that small repeated sequences are involved in genomic rearrangement (André *et al.* 1992). Considering the apparently irregular relationship of the flanking regions even in the two closely related species, wheat and rye, it is, therefore, not unpredictable that the organization of the flanking regions differs to some extent between rice and wheat/rye.

No plant species thus far investigated, except for rice, wheat, and rye, has

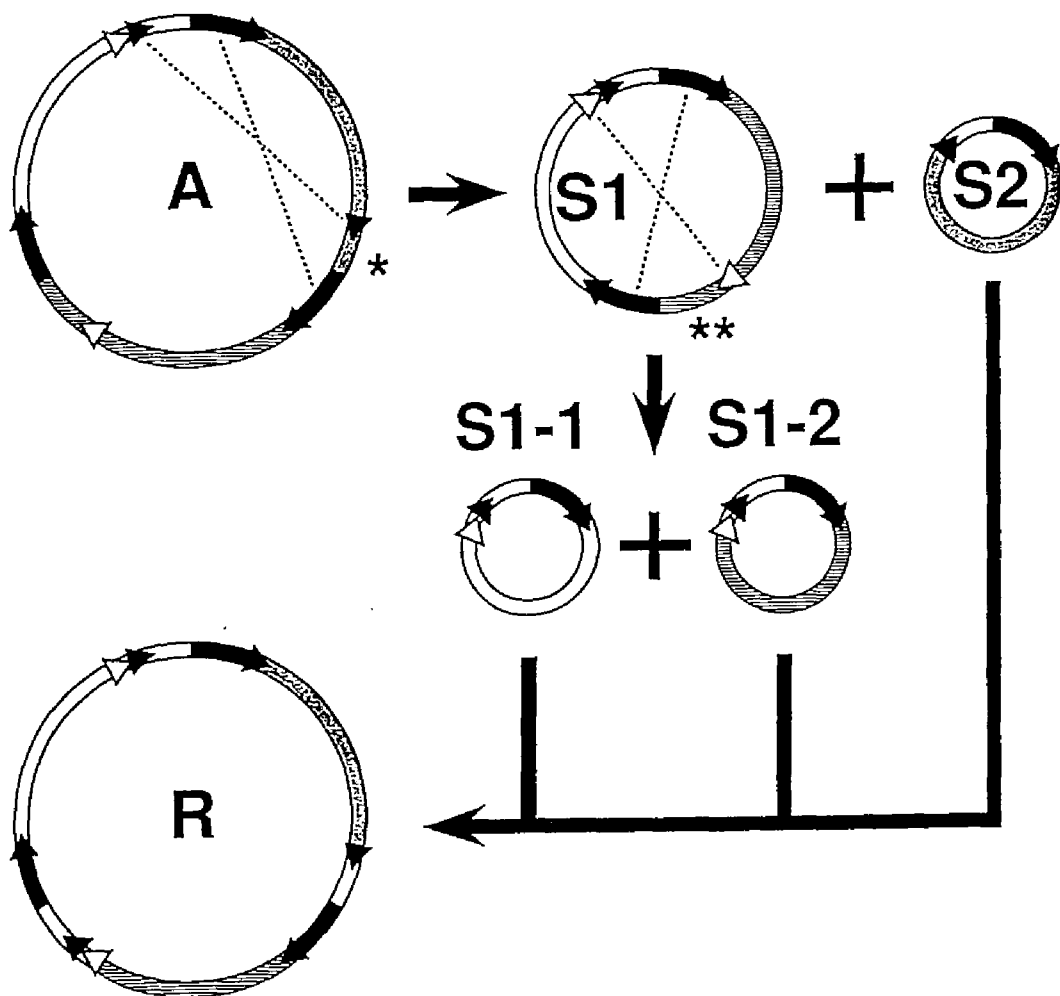


Fig. 3. Schematic illustration demonstrating how a rice-like genome configuration might be generated from a hypothetical ancestral genome common to rice, wheat, and rye. **A** is the ancestral genome and **R** is the rice-like genome, where filled boxes indicate direct repeats containing the *rrn18* and *rrn5* genes (18S/5S repeat) and triangles are short repeats at which rare homologous recombination occurs. Distinct regions separated by the 18S/5S repeats are shaded differently for clarity. **S1** and **S2** are the subgenomic molecules generated from genome **A** by homologous recombination between repeated sequences which are shown linked by dashed lines. Note that any subgenomic molecules containing the region marked by an asterisk are eliminated. Similarly, **S1-1** and **S1-2** are then generated from **S1**, and the region marked, this time by two asterisk, is eliminated. Genome **R** can then be organized from **S2**, **S1-1**, and **S1-2**.

been reported to possess multiple copies of the *rrn18/rrn5* genes. This indicates that the triplication of these genes is peculiar to rice, wheat, and rye. The mtDNA of the cultured A-58CMS cells was subjected to rearrangement during the process of callus induction, or the subsequent prolonged cell-culture, but preliminary experiments performed by the author indicated that the mitochondrial genome from shoots of A-58CMS actually possesses three copies of the genes. Iwahashi *et al.* (1992) found only two copies of the intact genes in the mitochondrial genome from a *Japonica*-type cultivar, Nipponbare (*O. sativa* L.), but the adjacent downstream region was revealed to exist in three copies. This is not inconsistent to the observations in A-58CMS, if independent evolution of a flanking region is taken into account. Further preliminary data revealed that *O. rufipogon* and *O. barthii*, which are closely related to the cultivated rice, *O. sativa*, possess at least two copies of the *rrn18/rrn5* genes and three copies of the downstream region, indicating that the *rrn18/rrn5* genes, or their neighbouring region, existed in a triplicated form in the ancestor of the genus *Oryza*.

There still remains another serious problem concerning the copy number of the *rrn18/rrn5* genes; has the maize mitochondrial genome never acquired three copies of the genes, or had it once acquired them but then lost two during the course of its evolution? If the latter possibility is the case, the copy number of the *rrn18/rrn5* genes can provide no information on the phylogeny of rice, wheat, and maize, but it is, at present, impossible to confirm this.

The location of rice in the phylogenetic tree of monocotyledonous plants has remained ambiguous despite the performance of diverse analyses based on a number of different aspects. The above discussion provides a rational scheme by interpreting the evidences in the simplest way possible; maize diverged first from the common ancestor of rice, wheat, and rye. To verify this scheme, however, much more evidence is required, including sequence configurations of the *rrn18* and *rrn5* genes in the other Graminae plants.

Summary

Chapter 1

The mitochondrial DNA (mtDNA) from cells cultured from a CMS line (A-58CMS) of rice (*Oryza sativa*) was cloned and its physical map constructed. Structural alteration occurred on the mitochondrial genome during cell culture. Detailed restriction analysis of cosmid clones possessing mtDNA fragments suggested either of the following possibilities: the master genome contains a 100 kb duplication (the genome size becomes 450 kb), or that a master circle is not present in the genome (the net structural complexity becomes 350 kb). To date, physical mapping of plant mitochondrial genomes has been illustrated in the form of a single circle, namely a master circle. No circular DNA molecule corresponding to a master circle has yet been proven, however. Representation of plant mitochondrial genomes and the possibility of a mitochondrial genome without a master circle are discussed.

Chapter 2

The nucleotide sequences of the rice mitochondrial 18S and 5S ribosomal RNA genes (*rrn18* and *rrn5*, respectively), along with their flanking regions, were determined. There are three copies of the *rrn18* gene in the mitochondrial genome of rice as in those of wheat and rye. Sequence analysis revealed that the triplicated region (5684 bp) of rice contains the intact coding sequences of the *rrn18* and *rrn5* genes, as well as the 5' flanking and the 3' flanking regions. Sequence comparison between rice and other plant species showed that the *rrn18* of rice has a large deletion in variable region 4, whereas the nucleotide sequences of the remaining coding regions are highly conserved.

Transcriptional analysis revealed that two transcription initiation sites (a possible consensus sequence is 'CATAGNRAA') exist for the rice *rrn18/rrn5* genes within the triplicated region, indicating equivalent contribution of the three copied rRNA genes to rRNA production.

Chapter 3

The phylogenetic relationship of rice, wheat and maize not yet have been clarified. The relationship among the flanking regions of the *rrn18* and *rrn5* genes in rice, wheat, rye, and maize was studied in detail, and the results, interpreted in the simplest way, implied that the phylogenetic divergence of maize was earlier than that of rice from wheat and rye.

References

- André C, Levy A, Walbot V (1992) Small repeated sequences and the structure of plant mitochondrial genomes. *Trends Genet* 8:128-132
- Attardi G (1985) Animal mitochondrial DNA: an extreme example of genetic economy. *Int Rev Cytol* 93:93-145
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K (1987) *Current Protocols in Molecular Biology*, Greene Publishing Associates and Willey-Interscience, Secaucus, NJ, p4.3.2
- Bendich AJ (1985) Plant mitochondrial DNA: unusual variation on a common theme. In: Hohn B, Dennis ES (eds) *Genetic flux in plants*. Springer, Wien, pp111-138
- Bendich AJ, Smith SB (1990) Moving pictures and pulsed-field gel electrophoresis show linear DNA molecules from chloroplasts and mitochondria. *Curr Genet* 17:421-425
- Biswas TK, Ticho B, Getz GS (1987) *In vitro* characterization of the yeast mitochondrial promoter using single-base substitution mutants. *J Biol Chem* 262:13690-13696
- Bland MM, Levings CS III, Matzinger DF (1986) The tobacco mitochondrial ATPase subunit 9 gene is closely linked to an open reading frame for a ribosomal protein. *Mol Gen Genet* 204:8-16
- Bonen L (1987) The mitochondrial S13 ribosomal protein gene is silent in wheat embryos and seedlings. *Nucl Acids Res* 15:10393-10404
- Brennicke A, Möller S, Blanz PA (1985) The 18S and 5S ribosomal RNA genes in *Oenothera* mitochondria: Sequence rearrangements in the 18S and 5S rRNA genes of higher plants. *Mol Gen Genet* 198:404-410
- Brown GG, Auchincloss AH, Covello PS, Gray MW, Menassa R, Singh M (1991) Characterization of transcription initiation sites on the soybean mitochondrial genome allows identification of a transcription-associated sequence motif. *Mol Gen Genet* 228:345-355
- Chétrit P, Mathieu C, Muller JP, Vedel F (1984) Physical and gene mapping of cauliflower (*Brassica oleracea*) mitochondrial DNA. *Curr Genet* 8:413-421
- Chowdhury MKU, Schaeffer GW, Smith RL, Matthews BF (1988) Molecular analysis of organelle DNA of different subspecies of rice and the genomic stability of mtDNA in tissue cultured cells of rice. *Theor Appl Genet* 76:533-539
- Christianson T, Rabinowitz M (1983) Identification of multiple transcriptional initiation sites on the yeast mitochondrial genome by *in vitro* capping with guanylyltransferase. *J Biol Chem*

- Clayton WD (1981) Evolution and distribution of grasses. *Ann Missouri Bot Gard* 68:5-14
- Coulthart MB, Huh GS, Gray MW (1990) Physical organization of the 18S and 5S ribosomal RNA genes in the mitochondrial genome of rye (*Secale cereale* L.). *Curr Genet* 17:339-346
- Coulthart MB, Spencer DF, Gray MW (1992) Comparative analysis of a recombining-repeat sequence family in the mitochondrial genomes of wheat (*Triticum aestivum* L.) and rye (*Secale cereale* L.). *Curr Genet*, in press
- Covello PS, Gray MW (1991) Sequence analysis of wheat mitochondrial transcripts capped in vitro: definitive identification of transcription initiation sites. *Curr Genet* 20:245-251
- Covello PS, Gray MW (1992) Silent mitochondrial and active nuclear genes for subunit 2 of cytochrome *c* oxidase (*cox2*) in soybean: evidence for RNA-mediated gene transfer. *EMBO J* 11:1815-1820
- Dale RM, McClure BA, Houchins JP (1985) A rapid single-stranded cloning strategy for producing a sequential series of overlapping clones for use in DNA sequencing: application to sequencing the corn mitochondrial 18S rDNA. *Plasmid* 13:31-40
- Dale RM, Mendu N, Ginsburg H, Kridl JC (1984) Sequence analysis of the maize mitochondrial 26S rRNA gene and flanking regions. *Plasmid* 11:141-150
- Dörfel P, Weihe A, Knösche R, Börner T (1989) Mitochondrial DNA of *Chenopodium album*(L.): a comparison of leaves and suspension cultures. *Curr Genet* 16:375-380
- Enomoto S, Ogihara Y, Tsunewaki K (1985) Studies on the origin of crop species by restriction endonuclease analysis of organellar DNA. I. Phylogenetic relationships among ten cereals revealed by the restriction fragment patterns of chloroplast DNA. *Jpn J Genet* 60:411-424
- Escote-Carlson LJ, Gabay-Laughnan S, and Laughnan JR (1990) Nuclear genotype affects mitochondrial genome organization of CMS-S maize. *Mol Gen Genet* 223:457-464
- Falconet D, Delorme S, Lejeune B, Sévignac M, Delcher E, Bazetoux S, Quétier F (1985) Wheat mitochondrial 26S ribosomal RNA gene has no intron and is present in multiple copies arising by recombination. *Curr Genet* 9:169-174
- Falconet D, Lejeune B, Quétier F, Gray MW (1984) Evidence for homologous recombination between repeated sequences containing 18S and 5S ribosomal RNA genes in wheat mitochondrial DNA. *EMBO J* 3:297-302
- Falconet D, Sévignac M, Quétier F (1988) Nucleotide sequence and determination of the extremities of the 26S ribosomal RNA gene in wheat mitochondria: evidence for sequence rearrangements in the ribosomal genes of higher plants. *Curr Genet* 13:75-82
- Fauron CM-R, Havlik M, Brettell RIS (1990) The mitochondrial genome organization of a maize fertile cmsT revertant line is generated through recombination between two sets of repeats. *Genetics* 124:423-428

- Fauron C, Havlik M, Lonsdale D, Nichols L (1989) Mitochondrial genome organization of the maize cytoplasmic male-sterile type T. *Mol Gen Genet* 216:395-401
- Folkerts O, Hanson MR (1989) Three copies of a single recombination repeat occur on the 443 kb mastercircle of the *Petunia hybrida* 3704 mitochondrial genome. *Nucl Acids Res* 17:7345-7357
- Folkerts O, Hanson MR (1991) The male sterility-associated *pcf* gene and the normal *atp9-1* gene in petunia are located on different mitochondrial DNA molecules. *Genetics* 129:885-895
- Fukuchi M, Shikanai T, Kossykh VG, Yamada Y (1991) Analysis of nuclear sequences homologous to the B4 plasmid-like DNA of rice mitochondria; evidence for sequence transfer from mitochondria to nuclei. *Curr Genet* 20:487-494
- Grabau EA (1985) Nucleotide sequence of the soybean mitochondrial 18S rRNA gene: evidence for a slow rate of divergence in the plant mitochondrial genome. *Plant Mol Biol* 5:119-124
- Gray MW, Hanic-Joyce PJ, Covello PS (1992) Transcription, processing and editing in plant mitochondria. *Annu Rev Plant Physiol Plant Mol Biol* 43:145-175
- Gualberto JM, Wintz H, Weil J-H, Grienemberger JM (1988) The genes coding for subunit 3 of NADH dehydrogenase and for ribosomal protein S12 are present in the wheat and maize mitochondrial genomes and are co-transcribed. *Mol Gen Genet* 215:118-127
- Hamby RK, Zimmer EA (1988) Ribosomal RNA sequences for inferring phylogeny within the grass family (*Poaceae*). *Plant Syst Evol* 160:29-37
- Hanic-Joyce PJ, Gray MW (1991) Accurate transcription of a plant mitochondrial gene in vitro. *Mol Cell Biol* 11:2035-2039
- Hanic-Joyce PJ, Spencer DF, Gray MW (1990) *In vitro* processing of transcripts containing novel tRNA-like sequences ('t-elements') encoded by wheat mitochondrial DNA. *Plant Mol Biol* 15:551-559
- Hiesel R, Brennicke A (1983) Cytochrome oxidase subunit II gene in mitochondria of *Oenothera* has no intron. *EMBO J* 2:2173-2178
- Hiesel R, Brennicke A (1985) Overlapping reading frames in *Oenothera* mitochondria. *FEBS Lett* 193:164-168
- Hiesel R, Schobel W, Schuster W, Brennicke A (1987) The cytochrome oxidase subunit I and subunit III genes in *Oenothera* mitochondria are transcribed from identical promoter sequences. *EMBO J* 6:29-34
- Houchins JP, Ginsburg H, Rohrbaugh M, Dale RM, Schardl CL, Hodge TP, Lonsdale DM (1986) DNA sequence analysis of a 5.27-kb direct repeat occurring adjacent to the regions of S-episome homology in maize mitochondria. *EMBO J* 5:2781-2788
- Hunt MD, Newton KJ (1991) The NCS3 mutation: genetic evidence for the expression of ribosomal protein genes in *Zea mays* mitochondria. *EMBO J* 10:1045-1052

- Ina Y (1992) The Manual for ODEn: Molecular Evolutionary Analysis System for DNA and Amino Acid Sequences, version 1.1.1, National Institute of Genetics, Mishima, Japan
- Iwahashi M, Nakazono M, Kanno A, Sugino K, Ishibashi T, Hirai A (1992) Genetic and physical maps and a clone bank of mitochondrial DNA from rice. *Theor Appl Genet* 84:275-279
- Joyce PBM, Spencer DF, Gray MW (1988) Multiple sequence rearrangements accompanying the duplication of a tRNA^{Pro} gene in wheat mitochondrial DNA. *Plant Mol Biol* 11:833-843
- Kadowaki K (1989) Ph.D thesis: Molecular biological studies on mitochondria of rice with male-sterile cytoplasm. Kyoto University, Japan
- Kadowaki K, Suzuki T, Kazama S (1989) Nucleotide sequence of the cytochrome oxidase subunit I gene from rice mitochondria. *Nucl Acids Res* 17:7519
- Kadowaki K, Suzuki T, Kazama S (1990) A chimeric gene containing the 5' portion of *atp6* is associated with cytoplasmic male-sterility of rice. *Mol Gen Genet* 224:10-16
- Kaleikau EK, André CP, Doshi B, Walbot V (1990) Sequence of the rice mitochondrial gene for apocytochrome b. *Nucl Acids Res* 18:372
- Kaleikau EK, André CP, Walbot V (1992) Structure and expression of the rice mitochondrial apocytochrome b gene (*cob-1*) and pseudogene (*cob-2*). *Curr Genet* 22:463-470
- Kawano S, Kuroiwa T (1985) Isolation and characterization of a membrane-DNA complex in the mitochondria of *Physarum polycephalum*. *Exp Cell Res* 161:460-472
- Kikuchi S, Takaiwa F, Oono K (1987) Variable copy number DNA sequences in rice. *Mol Gen Genet* 210:373-380
- Kuroiwa T (1982) Mitochondrial nuclei. *Int Rev Cytol* 75:1-59
- Levy AA, André CP, Walbot V (1991) Analysis of a 120-kilobase mitochondrial chromosome in maize. *Genetics* 128:417-424
- Lichtenstein C, Draper J (1985) Genetic engineering of plants. In: Glover DM (ed) DNA cloning: volume II: a practical approach, IRL Press, Oxford
- Lonsdale DM (1989) The plant mitochondrial genome. In: Stumpf PK, Conn EE (eds) The biochemistry of plants, vol 15. Academic Press, New York, pp 229-295
- Lonsdale DM, Brears T, Hodge TP, Melville SE, Rottmann WH (1988) The plant mitochondrial genome: homologous recombination as a mechanism for generating heterogeneity. *Phil Trans R Soc Lond B* 319:149-163
- Lonsdale DM, Hodge TP, Fauron CMR (1984) The physical map and organization of the mitochondrial genome from the fertile cytoplasm of maize. *Nucl Acids Res* 12:9249-9261
- Lonsdale DM, Hodge TP, Stoehr PJ (1986) Analysis of the genome structure of plant mitochondria. *Methods Enzymol* 118:453-470

- Mackenzie SA, Chase CD (1990) Fertility restoration is associated with loss of a portion of the mitochondrial genome in cytoplasmic male-sterile common bean. *Plant Cell* 2:905-912
- Mackenzie SA, Pring DR, Bassett MJ, Chase CD (1988) Mitochondrial DNA rearrangement associated with fertility restoration and cytoplasmic reversion to fertility in cytoplasmic male sterile *Phaseolus vulgaris* L. *Proc Natl Acad Sci USA* 85:2714-2717
- Maloney AP, Walbot V (1990) Structural analysis of mature and dicistronic transcripts from the 18S and 5S ribosomal RNA genes of maize mitochondria. *J Mol Biol* 213:633-649
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning: a laboratory manual, pp282-285, Cold Spring Harbor Laboratory Press, New York
- McNay JW, Chourey PS, Pring DR (1984) Molecular analysis of genomic stability of mitochondrial DNA in tissue cultured cells of maize. *Theor Appl Genet* 67:433-437
- Miyakawa I, Sando N, Kawano S, Nakamura S, Kuroiwa T (1987) Isolation of morphologically intact mitochondrial nucleoids from the yeast *Saccharomyces cerevisiae*. *J Cell Sci* 88:431-439
- Miyazawa S (1990) The Manual of the Flat Database and Sequence Analysis System for DNA and Proteins, version 1.3, National Institute of Genetics, Mishima, Japan
- Morikami A, Nakamura K (1987) Structure and expression of pea mitochondrial F₁ATPase α -subunit gene and its pseudogene involved in homologous recombination. *J Biochem* 101:967-976
- Mulligan RM, Lau GT, Walbot V (1988a) Numerous transcription initiation sites exist for the maize mitochondrial genes for subunit 9 of the ATP synthase and subunit 3 of cytochrome oxidase. *Proc Natl Acad Sci USA* 85:7998-8002
- Mulligan RM, Leon P, Walbot V (1991) Transcriptional and posttranscriptional regulation of maize mitochondrial gene expression. *Mol Cell Biol* 11:533-543
- Mulligan RM, Maloney AP, Walbot V (1988b) RNA processing and multiple transcription initiation sites result in transcript size heterogeneity in maize mitochondria. *Mol Gen Genet* 211:373-380
- Narayanan KK, André CP, Yang J, Walbot V (1992) Organization of a 117 kb circular mitochondrial chromosome in IR36 rice. *Curr Genet*, in press.
- Nei M (1989) Molecular Evolutionary Genetics, translated by Gojobori T & Saito N from the original English language edition (Columbia University Press, New York, USA), Baifu-kan, Japan
- Nishibayashi S, Kuroiwa T (1985) Division of mitochondrial nuclei in protozoa, a green alga and a higher plant. *Cytologia* 50:75-82
- Nugent JM, Palmer JD (1991) RNA-mediated transfer of the gene *coxII* from the mitochondrion to the nucleus during flowering plant evolution. *Cell* 66:473-481

- Nozato N, Oda K, Yamato K, Ohta E, Takemura M, Akashi K, Fukuzawa H, Ohshima K (1993) Cotranscriptional expression of mitochondrial genes for subunits of NADH dehydrogenase, *nad5*, *nad4*, *nad2* in *Marchantia polymorpha*. Mol Gen Genet, in press
- Oda K, Kohchi T, Ohshima K (1992a) Mitochondrial DNA of *Marchantia polymorpha* as a single circular form with no incorporation of foreign DNA. Biosci Biotech Biochem 56:132-135
- Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohshima K (1992b) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. J Mol Biol 223:1-7
- Ohshima K, Fukuzawa H, Kohchi T, Shirai H, Sano T, Sano S, Umesono K, Shiki Y, Takeuchi M, Chang Z, Aota S, Inokuchi H, Ozeki H (1986) Chloroplast gene organization deduced from complete sequence of liverwort *Marchantia polymorpha* chloroplast DNA. Nature 322:572-574
- Palmer JD, Herbon LA (1986) Tricircular mitochondrial genomes of *Brassica* and *Raphanus*: reversal of repeat configurations by inversion. Nucl Acids Res 14:9755-9764
- Palmer JD, Herbon LA (1987) Unicircular structure of the *Brassica hirta* mitochondrial genome. Curr Genet 11:565-570
- Palmer JD, Herbon LA (1988) Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. J Mol Evol 28:87-97
- Palmer JD, Shields CR (1984) Tripartite structure of the *Brassica campestris* mitochondrial genome. Nature 307:437-440
- Rode A, Hartmann C, Falconet D, Lejeune B, Quetier F, Benslimane A, Henry Y, de Buyser J (1987) Extensive mitochondrial DNA variation in somatic tissue cultures initiated from wheat immature embryos. Curr Genet 12:369-376
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a laboratory manual 2nd edition, pp7.79-7.83, Cold Spring Harbor Laboratory Press, New York
- Schinkel AH, Tabak HF (1989) Mitochondrial RNA polymerase: dual role in transcription and replication. Trends Genet 5:149-154
- Schuster W, Brennicke A (1987a) Nucleotide sequence of the ATPase 6 gene. Nucl Acids Res 15:9092
- Schuster W, Brennicke A (1987b) Plastid, nuclear and reverse transcriptase sequences in the mitochondrial genome of *Oenothera*: is genetic information transferred between organelles via RNA? EMBO J 6:2857-2863
- Scowcroft WR, Larkin PJ (1988) Somaclonal variation. In: Bock GB and Marsh J (eds) Applications of plant cell and tissue culture, pp21-35, John Wiley & Sons, Chichester
- Shikanai T, Yang ZQ, Yamada Y (1987) Properties of the circular plasmid-like DNA B1 from

mitochondria of cytoplasmic male-sterile rice. *Plant Cell Physiol* 28:1243-1251

- Shirzadegan M, Christey M, Earle ED, Palmer JD (1989) rearrangement, amplification, and assortment of mitochondrial DNA molecules in cultured cells of *Brassica campestris*. *Theor Appl Genet* 77:17-25
- Siculella L, Palmer JD (1988) Physical and gene organization of mitochondrial DNA in fertile and male sterile sunflower. CMS-associated alterations in structure and transcription of the *atpA* gene. *Nucl Acids Res* 16:3787-3799
- Small I, Suffolk R, Leaver CJ (1989) Evolution of plant mitochondrial genomes via substoichiometric intermediates. *Cell* 58:69-76
- Small ID, Isaac PG, Leaver CJ (1987) Stoichiometric differences in DNA molecules containing the *atpA* gene suggest mechanisms for the generation of mitochondrial genome diversity in maize. *EMBO J* 6:865-869
- Spencer DF, Bonen L, Gray MW (1981) Primary sequence of wheat mitochondrial 5S ribosomal ribonucleic acid: functional and evolutionary implications. *Biochem* 20:4022-4029
- Spencer DF, Schnare MN, Coulthart MB, Gray MW (1992) Sequence and organization of a 7.2 kb region of wheat mitochondrial DNA containing the large subunit (26S) rRNA gene. *Plant Mol Biol* 20:347-352
- Spencer DF, Schnare MN, Gray MW (1984) Pronounced structural similarities between the small subunit ribosomal RNA genes of wheat mitochondria and *Escherichia coli*. *Proc Natl Acad Sci USA* 81:493-497
- Stern DB, Palmer JD (1986) Tripartite mitochondrial genome of spinach: physical structure, mitochondrial gene mapping, and locations of transposed chloroplast DNA sequences. *Nucl Acids Res* 14:5651-5666
- Thomas CM (1992) Sugarbeet minicircular mitochondrial DNAs: high-resolution transcript mapping, transcript abundance and copy number determination. *Mol Gen Genet* 234:457-465
- Traynor PL, Levings CS III (1986) Transcription of the S-2 maize mitochondrial plasmid. *Plant Mol Biol* 7:255-263
- Wolfe KH, Gouy M, Yang Y-W, Sharp PM, Li W-H (1989) Date of the monocot-dicot divergence estimated from chloroplast DNA sequence data. *Proc Natl Acad Sci USA* 86:6201-6205
- Yamato K, Nozato N, Oda K, Ohta E, Takemura M, Akashi K, Ohyama K (1993) Occurrence and transcription of genes for *nad1*, *nad3*, *nad4L*, and *nad6* coding for NADH dehydrogenase subunits 1, 3, 4L, and 6 in liverwort mitochondria. *Curr Genet*, in press
- Yamato K, Ogura Y, Kanegae T, Yamada Y, Ohyama K (1992) Mitochondrial genome structure of rice suspension culture from cytoplasmic male-sterile line (A-58CMS): reappraisal of the master circle. *Theor Appl Genet* 83:279-288

List of Publications

- (1) Ohyama K, Kohchi T, Ogura Y, Oda K, Yamato K, Sano T, Yamada Y (1990) Organization and expression of chloroplast genome from a liverwort, *Marchantia polymorpha*. Bot Mag Tokyo Special Issue 2:145-158
- (2) Ohyama K, Ogura Y, Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Yamada Y (1991) Evolution of Organellar Genomes. in Evolution of Life (Osawa S & Honjo T eds, Springer-Verlag, Tokyo), pp187-198.
- (3) Ogura Y, Takemura M, Oda K, Yamato K, Ohta E, Fukuzawa H, Ohyama K (1992) Cloning and nucleotide sequence of a *frxC*-ORF469 gene cluster of *Synechocystis* PCC6803: conservation with liverwort chloroplast *frxC*-ORF465 and *nif* operon. Biosci Biotech Biochem 56:788-793
- (4) Yamato K, Ogura Y, Kanegae T, Yamada Y, Ohyama K (1992) Mitochondrial genome structure of rice suspension culture from cytoplasmic male sterile line (A-58CMS): reappraisal of the master circle. Theor Appl Genet 83:279-288
- (5) Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Gene organization deduced from the complete sequence of liverwort *Marchantia polymorpha* mitochondrial DNA: a primitive form of plant mitochondrial genome. J Mol Biol 223:17
- (6) Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Kanegae T, Ogura Y, Kohchi T, Ohyama K (1992) Complete nucleotide sequence of the mitochondrial DNA from a liverwort, *Marchantia polymorpha*. Plant Mol Biol Rep 10:105-163
- (7) Takemura M, Oda K, Yamato K, Ohta E, Nakamura Y, Nozato N, Akashi K, Ohyama K (1992) Gene clusters for ribosomal proteins in the mitochondrial genome of a Liverwort, *Marchantia polymorpha*. Nucl Acids Res 20:3199-3205

- (8) Oda K, Yamato K, Ohta E, Nakamura Y, Takemura M, Nozato N, Akashi K, Ohyama K (1992) Transfer RNA genes in the mitochondrial genome from a liverwort, *Marchantia polymorpha*: the absence of chloroplast-like tRNAs. Nucl Acids Res 20: 3773-3777
- (9) Nozato N, Oda K, Yamato K, Ohta E, Takemura M, Akashi K, Fukuzawa H, Ohyama K (1992) Cotranscriptional expression of mitochondrial genes for subunits of NADH dehydrogenase, *nad5*, *nad4*, *nad2* in *Marchantia polymorpha*. Mol Gen Genet, in press
- (10) Yamato K, Nozato N, Oda K, Ohta E, Takemura M, Akashi K, Ohyama K (1992) Occurrence and transcription of genes for *nad1*, *nad3*, *nad4L*, and *nad6* coding for NADH dehydrogenase subunits 1, 3, 4L, and 6 in liverwort mitochondria. Curr Genet, in press
- (11) Yamato K, Omata K, Nagai J, Ohyama K (1993) Structure, transcription, and evolution of three copies of 18S and 5S ribosomal RNA genes from rice mitochondria. in preparation

Acknowledgements

The author wishes to express his sincere gratitude to Professor Dr. Kanji Ohyama, Laboratory of Plant Molecular Biology, Faculty of Agriculture, Kyoto University, for his courteous guidance and continuous encouragement throughout the course of this study.

Special thanks must be given to Professor Dr. Yasuyuki Yamada, Associate Professor Dr. Fumihiko Sato, and Dr. Takashi Hashimoto, Laboratory of Molecular and Cellular Biology, Faculty of Agriculture, Kyoto University, for their thoughtful and kind encouragement.

This study could not have been accomplished without the substantial contributions of the author's colleagues in the laboratory, Dr. Yutaka Ogura, presently at the Research Institute for Bioresources, Okayama University; Mr. Takeshi Kanegae, presently at the Laboratory of Molecular and Cellular Biology; Mr. Kazuhiro Omata; Mr. Jun-ichi Nagai; Mr. Hisayoshi Sugawara, presently at Nikka Whisky, Co., Ltd.; and, Mr. Koh-ichi Uemura, Unitica, Co., Ltd., and the author is deeply grateful to them.

The author appreciates the valuable suggestions given to him by Professor Dr. Toshiro Kinoshita and Associate Professor Dr. Koh-ichi Mori, Hokkaido University; Dr. Koh-ichi Kadowaki, National Institute of Agrobiological Resources, Japan; Professor Dr. Koichiro Tsunewaki, Dr. Takashige Ishii, and Dr. Atsushi Ikeda, Laboratory of Genetics, Faculty of Agriculture, Kyoto University; Professor Dr. Michael W. Gray and Dr. Michael B. Coulthart, Department of Biochemistry, Dalhousie University, Canada; and Dr. Yasuo Ina, National Institute of Genetics, Japan. The author especially wishes to thank Dr. Toshiharu Shikanai, Nichirei, Co., Ltd., and Miss Masako Fukuchi, Suntory, Co., Ltd, for their technical advice.

The author gratefully acknowledges the criticisms, discussions, suggestions and considerate encouragements given to him by the members of the laboratory, in particular, Dr. Hideya Fukuzawa, Dr. Takayuki Kohchi, Dr. Kenji Oda, Mr. Eiji Ota, Miss Miho Takemura, Mr. Keun-Sik Lee, and Mr. Kin-ya Akashi.

Lastly, genuine thanks are due to my parents, sister, Uncle Shinzo, Aunt Nobuko, and friends who have affectionately supported the author.

March 1993

Katsuyuki Yamato
Katsuyuki Yamato